VOLUME 27 NUMBER 6 JUNE 2009
www.nature.com/naturebiotechnology

nature biotechnology

THE SCIENCE AND BUSINESS OF BIOTECHNOLOGY

Pichia pastoris sequence
Facilitating siRNA delivery
siRNA and endogenous miRNA competition

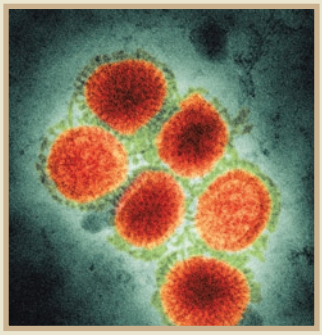
nature biotechnology

www.nature.com/naturebiotechnology

VOLUME 27 NUMBER 6 JUNE 2009

nature biotechnology

Scanning electron micrograph of *Pichia pastoris*. De Schutter *et al.* present the complete genomic sequence of this important protein expression system (p 561). Credit: Dennis Kunkel Microscopy, Inc.



Vaccine makers scramble to counter the emergence of a new flu strain, p 489

EDITORIAL

NEWS

BIOENTREPRENEUR

BUILDING A BUSINESS

OPINION AND COMMENT

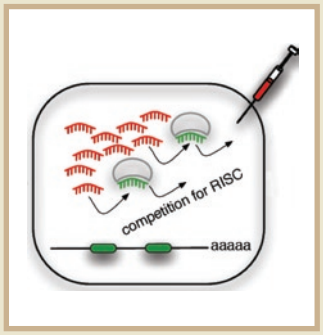
CORRESPONDENCE

Nature Biotechnology (ISSN 1087-0156) is published monthly by Nature Publishing Group, a trading name of Nature America Inc. located at 75 Varick Street, FI 9, New York, NY 10013-1917. Periodicals postage paid at New York, NY and additional mailing post offices. **Editorial Office:** 75 Varick Street, FI 9, New York, NY 10013-1917. Tel: (212) 726 9335, Fax: (212) 696 9753. **Annual subscription rates:** USA/Canada: US$250 (personal), US$3,060 (institution), US$4,050 (corporate institution). Canada add 7% GST #104911595RT001; Euro-zone: €202 (personal), €2,430 (institution), €3,488 (corporate institution); Rest of world (excluding China, Japan, Korea): £130 (personal), £1,570 (institution), £2,250 (corporate institution); Japan: Contact NPG Nature Asia-Pacific, Chiyoda Building, 2-37 Ichigayatamachi, Shinjuku-ku, Tokyo 162-0843. Tel: 81 (03) 3267 8751, Fax: 81 (03) 3267 8746. **POSTMASTER:** Send address changes to *Nature Biotechnology*, Subscriptions Department, 342 Broadway, PMB 301, New York, NY 10013-3910. **Authorization to photocopy** material for internal or personal use, or internal or personal use of specific clients, is granted by Nature Publishing Group to libraries and others registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided the relevant copyright fee is paid direct to CCC, 222 Rosewood Drive, Danvers, MA 01923, USA. Identification code for *Nature Biotechnology*: 1087-0156/04. **Back issues:** US$45, Canada add 7% for GST. CPC PUB AGREEMENT #40032744. Printed by Publishers Press, Inc., Lebanon Junction, KY, USA.

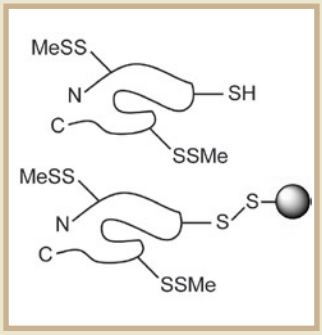
BPA WORLDWIDE

npg
nature publishing group

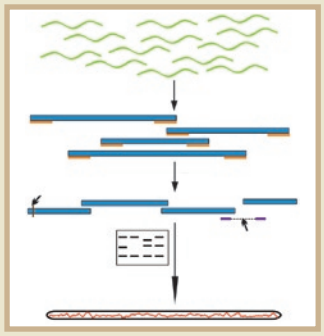
The future of forest biotech, p 519



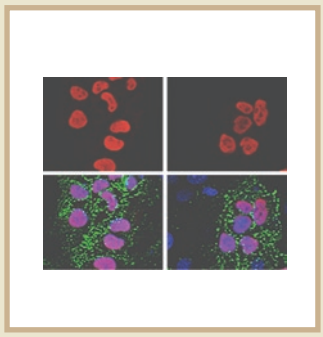
RISCy competition, p 549



Mass spec analysis of S-nitrosylation, p 557

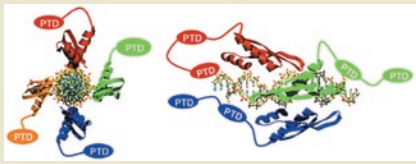


Pichia pastoris genome assembly, p 561



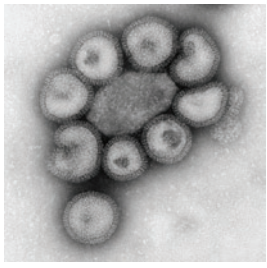
siRNA delivery, p 567

siRNA delivery to primary cells



RNA interference has become one of the most widely used techniques in basic biological research, drug target screening and target validation. Small interfering RNAs are progressing through the clinic as a potential therapeutic for certain diseases. Nevertheless, the delivery of siRNAs to primary cells with high efficiency and without cytotoxicity remains problematic. Eguchi *et al.* have addressed this problem by developing a fusion protein consisting of a peptide transduction domain and a double-stranded RNA binding domain. Peptide transduction domains can deliver macromolecules to the cytoplasm of most cells, whereas the double-stranded RNA binding domain masks the high negative charge of the RNA, which would otherwise increase the difficulty of shuttling the nucleic acid across cell membranes. Using this fusion protein, they deliver siRNAs to a range of difficult-to-transfect cells—including mouse T cells, human umbilical vein endothelial cells and human embryonic stem cells—with extremely high efficiency. They then use transcription profiling to show that the new delivery system induces no cytotoxicity, in contrast to commonly used lipid-based transfection reagents. The authors also demonstrate the *in vivo* potential of the fusion protein by knocking down luciferase expression in the nasal epithelium of transgenic mice. The availability of an efficient delivery system promises to extend the use of siRNA to many interesting cell types of high medical and biological importance. [**Letters, p. 567**] *ME*

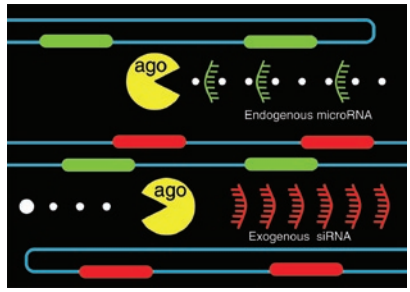
Influenza attenuation



Live attenuated viruses often make better vaccines than killed viruses, but generating the attenuated phenotype has long been a hit-or-miss process. Even when successful, repeated passaging of viruses in cell culture and selection for reduced fitness lead to viral strains that have disparate and unknown mutations. More recently, vaccine researchers have been developing methods to confer attenuation through genetic engineering. tenOever and colleagues have now applied such an approach to influenza A virus. By slightly altering the sequence of the viral nucleoprotein so that it contains target sites for the microRNA miR-93, the authors generate a virus that is attenuated in mice and potentially in humans, both of which express miR-93. Because chickens do not express this microRNA, the virus can still be grown to high titer in eggs. The authors propose that incorporation of microRNA target sites offers an alternative means of producing attenuated flu vaccines and might be usefully combined with the live attenuated vaccine FluMist—which is approved for people 2–49 years of age—to increase its safety. [**Letters, p. 572**] *KA*

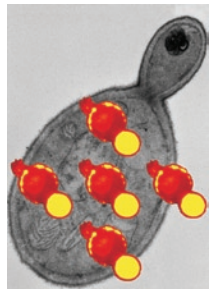
Written by Kathy Aschheim, Markus Elsner, Michael Francisco, Peter Hare, Craig Mak, & Lisa Melton

RISCy competition



Gene knockdown through RNA interference has been observed to generate unexpected side effects in the form of gene upregulation. To explain these effects, Khan *et al.* reanalyze data from 151 published studies in which cells were transfected with small RNAs (such as microRNA (miRNA), small interfering RNA (siRNA) or miRNA inhibitors) and the effects assayed by global gene expression profiling. Khan *et al.* look for evidence that the transfected siRNAs compete with endogenous miRNAs for cellular miRNA-processing machinery, such as the RISC complex. Notably, they find that endogenous miRNA targets—that is, genes with target sites recognized by miRNAs known to be expressed in the cell—are upregulated after transfection of an siRNA. Furthermore, based on patterns of unexpected transcript perturbation, Khan *et al.* are able to accurately predict the miRNAs expressed in a given cell type. This work demonstrates that effects consistent with a competition model are observable across a broad range of published studies that use varied cell types and siRNAs, suggesting that endogenous miRNAs should be considered in the design and interpretation of experiments using RNA interference. [**Analysis, p. 549**] *CM*

Pichia pastoris sequenced



The yeast *Pichia pastoris* is one of the most commonly used organisms for the production of proteins. Its importance only increased with the development of strains with fully humanized N-glycosylation. De Schutter *et al.* close an important gap in our understanding of this organism by presenting the complete 9.43 Mbp genomic sequence. Of the 5,313 protein-coding genes of *P. pastoris*, 75% have clear homologs in other species, but 71 of the gene families shared by the two closest sequenced relatives of *P. pastoris* (*P. stipitis* and *C. lusitaniae*) are absent from the *P. pastoris* genome. The authors identify many genes that are important for post-translational modifications, such as N- and O-glycosylation, the secretory pathway, and protein folding and degradation. These genes are likely targets for the engineering of optimized protein production

strains. The analysis of the codon usage bias and of endogenous signal sequences is also likely to be immediately useful in the design of optimized transgene constructs. What's more, the examination of the promoter regions of *P. pastoris* genes will allow devising highly efficient promotors for the expression of exogenous proteins. [**Letters, p. 561**] *ME*

Simpler analysis of *S*-nitrosylation

Although the ubiquitous and reversible modification of cysteines by nitric oxide can affect protein activity, localization and stability, its full roles in biology and disease remain poorly understood. As for most post-translational modifications, our insight into the functions of protein *S*-nitrosylation is limited primarily by technical bottlenecks. The biotin switch technique (BST) transformed this field almost a decade ago by enabling unbiased identification of *S*-nitrosylated proteins and their modification sites. But despite its eminence as the paramount methodology for studying *S*-nitrosylation *in vivo*, the multiple steps involved in the BST complicate its use for proteomic-scale analysis of protein *S*-nitrosothiols and limit its sensitivity. Stamler and colleagues modify the BST by changing the labeling reaction itself and coupling it directly to the purification of the derivatized protein, using a solid-phase reagent instead of a biotinylating reagent. At least in part as a consequence of the fewer acetone precipitation steps it involves, this approach (named SNO-RAC owing to the resin-assisted capture of *S*-nitrosothiols) is more sensitive than the BST, particularly for detecting SNO-proteins larger than ~100 kDa. The superior recovery of high-molecular-weight proteins enables identification of many previously unidentified *S*-nitrosylated proteins in several mammalian and bacterial cell types, as well as the specific sites of *S*-nitrosylation. By combining SNO-RAC with isobaric (iTRAQ) labeling to compare multiple samples in the same mass spectrometry experiment, the authors further demonstrate that SNO-proteins undergo dynamic denitrosylation on a global scale. Use of a thiol-reactive resin can be extended to other cysteine-based post-translational modifications, such as acylation. [**Brief Communications, p. 557**] *PH*

Patent roundup

After a seven-year legal battle, a US federal appeals court has ruled in favor of Eli Lilly over patent claims surrounding transcription factor NFκB. The decision overturns a prior decision that had disappointed many by granting Ariad broad rights relating to the NFκB pathway. [**News Analysis, p. 494**] *LM*

Germany is up in arms over a patent covering a marker-assisted test to breed meatier pigs. The patent, originally filed by Monsanto, is now owned by Newsham Choice Genetics. [**News in Brief, p. 496**] *LM*

A US Patent & Trademark Office filing deadline for patent term extension (PTE), which can add up to five years to a patent's life, is being frequently miscalculated by applicants. Dianna Goldenson explains that missing the PTE filing deadline is an incurable error, but missing it without getting caught opens a whole other can of worms. [**Patent article, p. 538**] *MF*

Recent patent applications in genome assays. [**New patents, p. 542**] *MF*

Next month in nature biotechnology

- Exploring drug synergy and selectivity
- siRNA delivery by nanocells
- Reproducibility and precision in targeted proteomics

EDITORIAL

nature biotechnology

The genome-assisted barnyard

In contrast to the slow translation of human genome information into medicine, animal genomics is likely to have a rapid and tangible impact on agriculture.

The immediate impact of the human genome sequence on human health and wellness has been rather underwhelming. Yes, sequence information from human and/or microbial genomes is immensely useful for understanding biology. But translating that into novel medicines and diagnostics is both complex and time consuming. Meanwhile, less anthropocentric genomic studies are forging ahead, with very little hype or fanfare. These livestock genomes are providing not only hugely valuable biological information but also immediate benefits to the way livestock breeders go about their business.

The latest milestone in animal genomics is the assembled sequence of *Bos taurus*, domestic cattle (*Science* **324**, 522–528, 2009; http://www.biomedcentral.com/series/bovine), and an analysis of >37,000 single nucleotide polymorphisms (SNPs) in 497 cattle from 19 geographically and biologically diverse breeds (*Science* **324**, 528–532, 2009). The bovine genome follows the chicken sequence (*Nature* **432**, 695–716, 2004) and precedes those of the pig and sheep, which are slated for release later this year and sometime in the next couple of years, respectively. All of these livestock sequences provide insights into gene function, evolution and the origins of different breeds. But they also provide something that simply cannot be exploited in humans.

Where human genomics does (somewhat rarely) shed light on human characteristics, the genetic information can be used only indirectly. It is considered immoral to breed out 'undesirable' traits in humans. Consequently, pinpointing a disease-associated allele in a human may improve our understanding of the condition and suggest a drug target or a plausible diagnostic agent, but its utility only emerges when (or if) the product development processes successfully run their course.

The case could not be more different for agriculturally important animals. For all types of livestock, genomic information is directly aligned with the means for improvement. First identify the gene(s); then breed (or splice) them in. In fact, it is not even necessary to identify the gene(s) or to understand the molecular processes behind milk yield or meat quality or the fecundity of sheep. All that is really needed is to be able to correlate the desired quality with the genetic variant and then to set up an appropriate breeding program.

Thus, working from the draft bovine sequence assembly, researchers led by Curt Van Tassell of the Agricultural Research Service at the US Department of Agriculture (USDA) teamed up with Illumina to create a BeadChip containing >50,000 bovine SNPs. This was commercially launched in January 2008, and by July last year, the dairy industry was using it to direct its breeding approach. Similarly, with <50% of the sheep sequence available, an Illumina Ovine SNP50 BeadChip released in January is already being used to analyze a population of 20,000 sheep in New Zealand and elsewhere. Part of the reason that SNP discovery can proceed so rapidly is the increasing use of reduced representation libraries and next-generation sequencing, which provide not only SNP positions but also concurrent estimates of minor allele frequencies (*Nat. Methods* **5**, 247–252, 2008), all at a cost of about $0.50 a SNP.

Another key difference from the human situation is that animal genomic information is not data in a vacuum. There are large databases on breeding and progeny stock for several of the main types of livestock animal, data that can be correlated with genetic data. Thus, for most animal genomes, a much bigger piece of the data jigsaw is already in place when genome sequence becomes available.

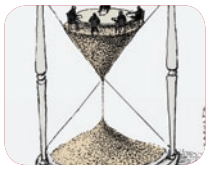
For example, New Zealand's sheep database was started in 1968 and now contains data for over 6.5 million animals on key characteristics, such as lamb survival, the number of lambs born, their birth weight, growth rates, disease resistance and meat yield and quality. For dairy cattle, the USDA has data from breed registry societies going back to 1960—over 60 million milk records covering more than ten generations of animals. The pork, beef and chicken records are less comprehensive, but even so the cataloging can go back at least as far as the great-grandparents of the current generation.

Centuries of directed improvement means that domestic animal variation is narrow, even between countries or climates. This means a genomic study of one cow or chicken is often directly relevant to other domesticated strains. Furthermore, the inbred nature of food animals makes the application of trait-associated markers more straightforward, especially when one breed predominates. This is one reason why the uptake of marker-assisted breeding has been relatively rapid in the US dairy industry, where >90% of cows are Holsteins.

The fact that at least three companies now market gene tests to breeders for economically important traits suggests demand for marker-assisted breeding exists. In March 2008, Pfizer Animal Health launched a new animal genetics division to market gene tests for quality grade, tenderness and feed efficiency. Elsewhere, Merial (Igenity) and Cargill (Metamorphix) have also put gene tests on the market for similar traits. Pricing competitively may be key to success in marker-assisted breeding and the market dynamics will almost certainly vary depending on the application of the approach. A $100 genetic test that is feasible for the sire of a dairy herd (large number of offspring and continuing benefit of raised milk yields) might be prohibitive for a beef herd (fewer animals per sire and a one-off product). Concerns are also emerging over the ability of companies to acquire intellectual property to protect the genes in their tests (see p. 496).

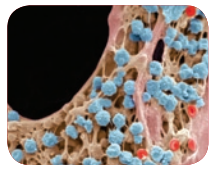
Nevertheless, the beauty of marker-assisted breeding is that it is likely to be less disconcerting to both the public and regulators than transgenic or cloned animals. Animal genomics may just be the accelerant that animal biotech needs to grab the spotlight from its more showy but often unproductive human health counterpart. Certainly, if marker-assisted breeding lives up to its early promise, it could change the face of animal health, welfare and productivity.

IN this section



Biotech companies wrestle with hostile investors p492

High-stakes tussle over NFκB patent claims sees Lilly defeat Ariad p494



Players in the cardiac cell therapy field wecome new guidelines p496

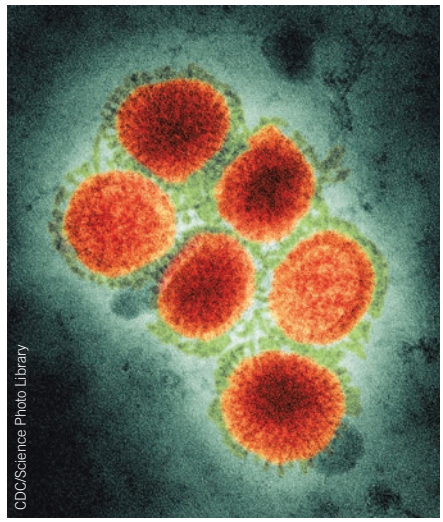
Flu vaccine makers upgrade technology—and pray for time

The onset of winter in the southern hemisphere could determine whether the novel, swine-derived influenza A (H1N1) strain detected in Mexico in April will evolve into a full-fledged pandemic virus. Health authorities around the globe are by no means fully equipped to cope with a severe influenza pandemic at this stage, given the lead times and capacity constraints associated with current vaccine production processes—as was the case in 2004, when a dangerous new avian strain (H5N1) emerged in Asia. In the past five years, however, public sector laboratories and biotech companies have made considerable progress in developing modern alternatives to the cumbersome, egg-based manufacturing process that large flu vaccine producers have relied on for decades.

One example is OptaFlu, made by Basel-based Novartis, the first product based on a more flexible mammalian cell culture process, which has recently become available in Europe. And in the past two years, according to a report published in February by New York–based consultants Oliver Wyman, potential global manufacturing capacity for pandemic vaccines has increased by 300% (for avian H5N1 strains at least) because of process improvements and dose-sparing strategies adopted by producers. Nevertheless, it will take several more years for innovative vaccines—based on recombinant approaches involving fusion proteins, DNA sequences or virus-like particles (VLPs)—to be available at the scale required to cope with a major pandemic.

As *Nature Biotechnology* went to press, the Geneva, Switzerland–based World Health Organization (WHO) had not altered its recommendations to the dozen or so large and small vaccine producers that are licensed to make seasonal vaccines based on hemagglutinin (HA) and neuraminidase (NA) viral antigens. "Unless we receive the order to produce H1N1 vaccine as a priority we will not be producing it," says Albert Garcia, spokesman for Sanofi Pasteur, the vaccines arm of Paris-based Sanofi Aventis, which produced about 170 million of the world's total supply of around 400 million seasonal flu vaccines last year. In May, the company received an FDA license for a second production facility at its Swiftwater, Pennsylvania, location, which will add another 100 million doses when fully up and running.

Production of seasonal, trivalent vaccine for the northern hemisphere's 2009–2010 flu season, based on WHO recommendations issued in February on the three most likely circulating strains, is nearing completion. A decision to move to a monovalent, pandemic vaccine would be based on a combination of the epidemiology



CDC/Science Photo Library

Virions from the H1N1 outbreak in April 2009. Universal vaccines based on highly conserved antigens could protect against this and multiple other influenza strains, but these vaccines are some way away.

Table 1 Selected pandemic influenza vaccines in development

Company (location)	Product	Status
GSK	Prepandrix, AS03 (oil in water), adjuvanted/prepandemic H5N1 vaccine	Dossier preregistered
Omninvest (Budapest, Hungary)	H5N1 whole-virus adjuvanted vaccine	Registered (in Hungary)
Novartis	Focetria, M59-adjuvanted/H5N1 prepandemic vaccine	Dossier preregistered
Baxter International (Deerfield, Illinois) Dynport Vaccine (El Segundo, California)	H5N1 inactivated whole virion/cell culture–derived inactivated vaccine	Phase 3
CSL (King of Prussia, Pennsylvania)	Inactivated H5N1 vaccine	Phase 2
Novavax (Rockville, Maryland)	Insect cell–derived, H5N1 virus-like particle based on HA, NA, M1 antigens	Phase 2
Sanofi Pasteur	H5N1 inactivated, split-virion vaccine	Phase 2
Vical (San Diego)	IPT1-101, DNA vaccine encoding NP, M2e and HA with cationic lipid adjuvant Vaxfectin	Phase 1
Avir Green Hills Biotechnology (Vienna)	FluVacc, live attenuated intranasal pandemic vaccine	Phase 1
Generex Biotechnology (Toronto)	Synthetic H5N1 vaccine, based on modified H5 peptides	Phase 1
AmVac (Zug, Switzerland) National Health Research Institutes (Zhunan, Taiwan)	MAPL-2 adjuvanted H5N1 intranasal vaccine	Preclinical
Inviragen (Fort Collins, Colorado)	Recombinant, intranasal MVA-based H5N1 vaccine	Preclinical
Medicago	Plant-cell-derived H5N1 virus-like particle based on HA antigen	Preclinical
MedImmune (Gaithersberg, Maryland) NIH (Rockville, Maryland)	16 subtype strains of live attenuated pandemic vaccine	Research
Nasvax (Ness-Ziona, Israel)	Ceramide carbamoyl spermine plus cholesterol-adjuvanted H5N1 vaccine	Research

IN brief

Taiwan builds biotech runway

ITRI

The BEL labs ensconced in Taiwan's ITRI.

Taiwan's government has announced the launch of a $1.76 billion venture capital fund as part of a comprehensive 'biotechnology takeoff package' aimed at putting the country on Asia's biotech map. The National Development Fund will have a 40% stake in the venture with the private sector contributing the rest. "We have a lot of early-stage discovery [in Taiwan] but a mechanism to commercialize it has been lacking," says Chong-Chou Lee, director of the biotech office in the government's science and technology advisory group. "We need this package to [bridge the gap between discovery and] clinical trials and then think about partners for tech transfer." Taiwan urgently needs an innovation-based biotech business to replace its manufacturing-based information and communications technology (ICT) industry. Taiwan's pool of medical professionals and well-equipped facilities have so far attracted companies such as Novartis of Basel and London-based GlaxoSmithKline, who have set up clinical R&D centers in the capital. But Taiwan's limited domestic market has meant that growth in the biopharma sector has been modest. With the newly launched venture capital fund, any biotech project with commercial potential, based in Taiwan, stands a good chance of being supported, provided the products eventually make it to the global market. "Innovation in biotechnology here is growing, but the challenge is to connect the local with the global," says Chung-Cheng Liu, general director of Biomedical Engineering Research Laboratories (BEL) in Taipei, the largest nonprofit R&D organization in Taiwan and part of the Industrial Technology Research Institute (ITRI). The fund will support the drive towards commercialization with measures aimed at strengthening the country's infrastructure by, for instance, setting up preclinical testing labs and establishing the Taiwan Food and Drug Administration to bring the regulatory environment to international standards. Biotech incubators will be created within existing science parks, positioning biopharmaceutical companies near the Academia Sinica (the Taiwanese National Institutes of Health) in Taipei, medical device firms close to BEL in the Hsinchu area and agricultural biotech at a site in south Taiwan. The venture fund is looking to boost the number of licensing deals. For example, the German firm Boehringer Ingelheim has recently signed an agreement to develop Taipei-based AbGenomics' novel monoclonal antibody to treat autoimmune diseases, and BEL and the National Taiwan University Hospital have inked a deal with Exactech, an orthopedics company located in Gainesville, Florida, to use a cartilage repair platform developed locally. *Susan Aldridge*

Table 2 Universal influenza vaccines in development

Developer	Product	Status
BiondVax (Ness Ziona, Israel)	HA, NP, M1 epitopes fused to bacterial flagellin protein	Phase 1
Sanofi Pasteur	Acam-Flu-ATM	Phase 1
Merck (Whitehouse Station, New Jersey)	IMX-adjuvanted bivalent influenza peptide conjugate vaccine	Phase 1
VaxInnate	Cell culture–derived recombinant M2e antigen fused to bacterial flagellin	Phase 1
Dynavax Technologies	CpG-adjuvanted recombinant protein containing nucleoprotein plus eight M2e repeats	Preclinical
FluGen (Madison, Wisconsin)	M2, live, attenuated H5N1 vaccine	Preclinical
Juvaris Biotherapeutics (Burlingame, California)	Adjuvanted vaccine for influenza A and B strains	Preclinical

and the severity of the virus, which, so far, has largely resulted in relatively mild illness except in Mexico, which for unexplained reasons, has experienced far more deaths than any other country.

In the meantime, Sanofi Pasteur—and other large influenza vaccine makers, such as GlaxoSmithKline (GSK), of London, Novartis and Baxter of Deerfield, Illinois—are gearing up to be ready to produce pandemic vaccine using seed strains propagated at WHO-designated labs. GSK estimates it will take 4 to 6 months to generate the vaccine. In mid-May, the UK Department of Health placed an order of up to 90 million doses of pandemic vaccine from GSK and Baxter, with the former company indicating that the governments of France, Belgium and Finland also intend to purchase 50 million, 12.6 million and 5.3 million doses, respectively.

Many biotech firms have also been developing novel flu vaccines (**Table 1**), based on alternative production methods. They have begun work on prototype H1N1 vaccines to demonstrate their capabilities—not least to potential large pharma partners, which, for the most part, have been slow to embrace novelty in this market. "There's a huge inertia in the system," says Alan Shaw, CEO of one such firm, Cranbury, New Jersey–based VaxInnate. That's because existing producers have invested large amounts of time and cash in optimizing their current production processes, which, though far from perfect, are cheap and usually (though not always) reliable.

VaxInnate, which closed a $30 million financing round in May, is developing both seasonal and pandemic vaccines, based on recombinant proteins that can be rapidly produced in high volumes in bacterial expression systems. A 1,000-liter fermentation process, Shaw says, could produce around 400 million doses in a matter of months. In each instance, VaxInnate is combining a viral antigen with an immunostimulatory bacterial flagellin protein, which binds Toll-like receptor (TLR)5 and triggers both an innate immune response and a more efficient adaptive response against the virus. Its seasonal vaccine incorporates a hemagglutinin (HA) antigen, whereas

Box 1 The near-term solution: antivirals

The two available flu drugs—the NA inhibitors Tamiflu (oseltamivir), developed by Basel-based Roche and Gilead Sciences, of Foster City, California, and Relenza (zanamivir), developed by GSK and Biota, of Notting Hill, Australia—appear to be effective against the current H1N1 strain (even though high levels of resistance to Tamiflu were reported in circulating seasonal strains in 2008–2009).

Elsewhere, BioCryst Pharmaceuticals, of Birmingham, Alabama, hopes to provide a third NA inhibitor, peramivir, which it is now positioning as an intravenous treatment for hospitalized patients, after a recent trial in patients receiving the drug by intramuscular injection failed to demonstrate efficacy. The company has sought an Emergency Use Authorization pending completion of its clinical development program. The drug's first approval could come in Japan, however, where its partner Shionogi, of Osaka, is running a pivotal trial. "They expect to be able to file their NDA [new drug application] in the current year," says BioCryst chief medical officer Bill Sheridan.

Other drug developers working on alternative technologies, including RNA interference–and antibody-based therapeutics, are at earlier stages of development. Pulmatrix, of Lexington, Massachusetts, is developing an aerosol formulation of undisclosed cations, which appears to have physical and biological effects that can treat and protect against infection. "It works preclinically—independent of the pathogen, independent of the strain and independent of the species," says CEO Robert Connelly. The company has completed a phase 1 study and a phase 1b study will start this summer. *CS*

its pandemic vaccine is based on a more highly conserved—but less immunogenic—antigen, the extracellular domain of the M2 viral matrix protein (M2e). "The real hurdle here is M2 has never been shown to protect humans against disease—it works well in mice," says Shaw.

Universal vaccines, based on highly conserved viral antigens, such as M2e, could provide multi-year protection against multiple influenza strains (**Table 2**). They could be stockpiled in advance allowing vaccine makers to get off the annual reformulation treadmill needed to keep up with the HA and NA antigens' mutability. Other recent work has suggested that a concealed hydrophobic pocket in the conserved stem region of HA might also be a conserved epitope suitable for vaccine development (*Nat. Struct. Mol. Biol.* **16**, 265–273, 2009; *Science*, published online, doi: 10.1126/science.1171491, February 26, 2009).

Several universal vaccines have already entered the clinic, but progress has been slow. "I believe one of the reasons these things have not moved very quickly is the results have not been spectacular," says Dino Dina, CEO of Dynavax Technologies, of Berkeley, California. Next year, Dynavax aims to start a clinical trial of another candidate universal vaccine, a recombinant protein comprising two conserved viral antigens, nucleoprotein (NP) and M2e, fused to an immunostimulatory sequence that acts as a TLR9 agonist. "Nucleoprotein generates immunity during natural infection, but it's only present in trace amounts in conventional vaccines," says Dina. The protein, he says, elicits a cytotoxic T-cell response, which could help to reduce viral spread and transmission. "In a pandemic kind of setting that would be a very valuable feature."

But others see universal vaccines as a long-term bet. "Our biggest concern is that the regulatory pathway for universal vaccines is not clear," says Rahul Singhvi, CEO of Rockville, Maryland–based Novavax. The firm uses VLP technology to develop vaccines based on HA and NA and the structural protein M1. A baculovirus vector expressed in an insect cell culture system produces particles, which closely resembles the native virus. "To the immune system it appears like there's a natural infection at the site of immunization," says Singhvi. This approach, he says, would enable large-scale manufacturing within around 12 weeks of a pandemic strain being characterized.

The company is also offering, in conjunction with GE Healthcare, a subsidiary of Fairfield, Connecticut–based GE, a low-cost, portable, disposable manufacturing system for pandemic vaccines. "You can do this in low-infrastructure environments," Singhvi says.

Quebec-based Medicago is also harnessing VLP technology, but in a radically different setting. The company has developed a transient gene expression system in the plant species *Nicotiana benthamiana*, a close relative of the tobacco plant, which can produce VLPs comprising the viral HA antigen only. It relies on an *Agrobacterium* plasmid to deliver the construct to the plant cells. Frederic Ors, Medicago's vice president of business development, says the purified VLPs are highly immunogenic, and the production process is also relatively low cost. "All you need is a greenhouse," says Ors. "The biomass production is cheap, even in comparison to eggs."

It will be several years yet, however, before any of these innovations—and others in development at competitor firms—will be ready for commercial rollout. In the meantime, drug therapy will remain a vital frontline defense against a pandemic (**Box 1**).

At this point, it is not yet clear whether the current pandemic alert will escalate further or will peter out, as recent avian flu epidemics have done. What is certain is that a vaccine for swine-derived H1N1 lies several months away.

Cormac Sheridan *Dublin*

Profiting from pandemics

It's most likely that established vaccine developers, such as London-based GlaxoSmithKline and Sanofi Aventis of Paris, would pump out stockpiles of any pandemic flu vaccine, but it is the small biotechs that literally rise and fall with the world's pandemic concerns. Note Birmingham, Alabama–based BioCryst, developer of the clinical stage neuraminidase inhibitor peramivir, for influenza. The firm received a 90% stock boost to $3.29 on April 27, after the H1N1 influenza (swine flu) grabbed headlines. And in London, Lipoxen on April 30 announced positive preclinical results for the delivery of an enhanced influenza vaccine, adding that the technology should also work against the new swine flu strain. Investors boosted Lipoxen's share price from £6.62 ($10.11) to £21.75 ($33.23). Also consider Rockville, Maryland–based vaccine developer Novavax. The company's stock slowly lost ground this year, dropping from $2 per share to around 85 cents in mid-April. But when swine flu became the topic of conversation, Novavax's shares jumped more than 200% to $2.55 over two sessions. Similarly, in 2005, when the flu was avian rather than swine, Novavax's shares traded at less than a dollar for most of that summer. However, in the fall, when the company's avian flu vaccine, manufactured using their virus-like particle technology, performed well in animal models, Novavax's stock jumped to close as high as $5.53. ***Brady Huggett***

Boardroom tensions rise as investors push for liquidation

When company management find themselves at odds with investors, things can get ugly very quickly. In March, Alameda, California–based Avigen had to defend itself against its biggest shareholder, the Biotechnology Value Fund, when the fund pressed the struggling company to cease operations and hand over the remaining cash. As the funding drought continues, concerns are growing that more and more public biotech companies may find themselves at the mercy of investors.

A certain amount of company attrition is natural. "Companies get caught in a downward spiral," says Corey Davis, senior analyst at Natixis Bleichroeder in New York, "The stock price drops and, if you can't raise the needed cash, that drops the price even more." But the credit crunch has caused many more biotechs to find themselves in the precarious position of being valued at less than their cash on hand.

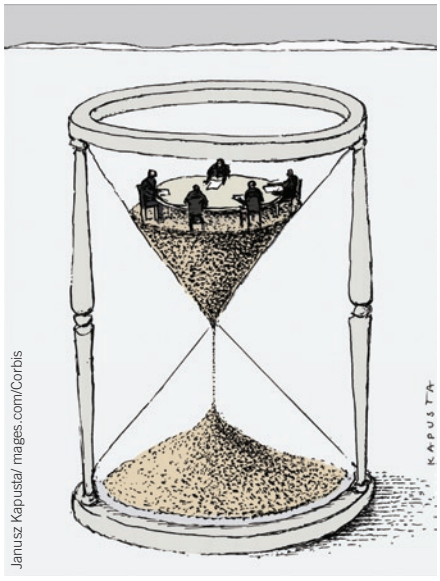
At the end of April, the investment bank Rodman & Renshaw in New York estimated that 49 biotech companies were trading below their cash values (**Table 1**). "Small-cap speculative investments have been hit hard because investors have been unwilling to take the risk," remarks Simos Simeonidis, senior biotech analyst at Rodman & Renshaw. "The value of many assets has decreased 50–70% on no real data." He says that companies are trading lower than their cash on hand because investors don't see much value in the companies' assets and believe that the companies will continue burning through the cash. "The companies are being valued at the level of cash they'll have 12 to 18 months from now."

According to Simeonidis, specialized investors are losing patience. "The core investors have leverage that allows them to push to liquidate. They don't want to wait five plus years for what they would refer to as a science experiment." In better times, when investors are more optimistic, the market value might be closer to the intrinsic value of the company, Davis explains. But with credit unavailable, investors aren't willing to take the risk. "It all comes back to money," says Davis. "If you can't raise money, you can't realize the return that's tied up in the assets."

Table 1 Biotech companies trading near cash levels

Market cap value	Number of companies
Below cash	49
Below net cash	32
Between one and two times cash	57
Between one and two times net cash	40

Source: Rodman & Renshaw



Janusz Kapusta/ mages.com/Corbis

Biotech companies trading below their cash value are running out of time.

Davis thinks the slumping stock market could be causing minority investors to be more willing to side with the activist than they were in the past. "Certain activists have the ear of other investors because the investors have lost so much," Davis says. "Investors always want to maximize value. Sometimes that means selling the company or shutting down. Because management never wants to sell at the bottom, activism is required to get things done."

The difference between intrinsic and market value is what led Avigen into a fight with its shareholders. Last fall, AV650, a controlled release form of tolperisone, the company's treatment for spasticity associated with multiple sclerosis, failed its phase 2b clinical trial causing shares to sink. Avigen has since won a proxy fight against Biotechnology Value Fund, a large shareholder based in New York, which wanted to purchase the company outright, but Avigen still plans on liquidating its assets and shutting down operations.

Investors don't always get it right. In February, Tang Capital Partners of San Diego proposed a resolution to request Rockville, Maryland–based Vanda Pharmaceuticals' board to "promptly take all necessary action to swiftly and orderly liquidate the Company's remaining assets and return all remaining capital to the Company's stockholders," according to documents filed with the Securities and Exchange Commission. Shares had fallen considerably after the company received a not approvable letter for its schizophrenia treatment, Fanapt, a sustained-release formulation of iloperidone, from the US Food and Drug Administration (FDA) in July of 2008. Management fought back, and in early May they were vindicated when the FDA approved Fanapt. The stock price soared and Tang Capital ended its proxy contest.

To fend off such hostile shareholder activity, management teams have instituted 'poison pills', giving all shareholders the right to acquire additional shares below market value when a hostile bidder tries to take over a company. The plans are able to stop hostile takeover attempts because existing shareholders exercising their right to buy additional shares dilute the hostile bidder's ownership of the company.

The problem is that poison pills may not hold off the bidder forever. But that's not the point, says Jeremy Grushcow, a US corporate lawyer practicing at Ogilvy Renault in Toronto. "Most companies being targeted by a hostile bidder that have a plan in place eventually get acquired. The purpose of a poison pill is often to gain leverage and force an acquirer to the table." Negotiations can result in terms that are more favorable for management, and may also help buy time for a second bidder to make a better takeover offer, says Grushcow.

Rather than negotiating, hostile investors can get rid of a poison pill by winning a proxy fight and gaining control of the board. Once investors take control, they can usually remove or neutralize the plan put in place by the company. "Companies fight that by staggering their boards so that board members are elected at different times," says Grushcow.

But gaining control of a board through multiple rounds of elections takes time, and investors often try to counter this delay by moving up a company's annual meeting. For instance, one of the proposed amendments by Tang Capital was to require Vanda's annual meeting be held on April 30

of each year. Companies often take the opposite position, delaying the annual meeting to buy more time, says Grushcow.

Another tactic called a 'poison put' is designed to protect bond holders by allowing creditors to call in a loan if there is a change in control of the company. The 'puts' are no longer just protecting creditors though; they're also helping management keep hold of the company, says Grushcow. "In an ordinary credit environment, companies could just refinance their debt, but now the put has become a threat rather than just an obstacle," he says.

Private companies are having a much easier time dealing with their investors. Although deal terms may be tougher for startups, there's still capital available from venture funds—at least for later-stage companies and "as long as companies have been successful at hitting milestones," explains Mark Lupa, partner at High Country Venture/Tango in Boulder, Colorado.

The initial public offering market is closed for now, but investors are still able to exit through acquisitions. The process may take longer, so established private companies are hunkering down and conserving cash, explains Chris Christoffersen, partner at Boulder-based Morgenthaler Ventures. He says that investors are also "putting more reserve funds into the deals to make sure companies get to the next round."

When companies don't hit their milestones, it's often easier to shut down or sell private companies because investors already have a seat on the board. "Things are always simpler for private companies. Information is much more available and it's easier to do things—whether it's building or taking apart," says Lupa.

Bruce Booth, partner at Atlas Venture in Massachusetts, agrees. "In private companies, it's generally a shared view that the prospects for a company's programs have sufficiently deteriorated to make raising new capital unattractive, if not impossible. Most of the time, significant efforts are made by both board and management to explore a full range of strategic alternatives—sale, merger or recapitalization, for example—before a shutdown."

It remains to be seen where the money that investors receive after companies are closed down will be reinvested. "It's going to go where the easiest money will be made. If the model works, investors will continue to invest in biotech," says Davis. But he cautions, "The jury is still out on whether it's working or not."

Brian Orelli *San Diego*

Still strapped for cash

The beginning of April marked the end for Irving, Texas–based DelSite. Although the company had FDA clearance for a phase 1 trial of its GelVac nasal powder H5N1 influenza vaccine, it had spent several months fruitlessly seeking funding to pay for the work. It found none, so on April 2, it announced its filing for bankruptcy protection, setting itself up for liquidation. Equity holders are expected to get nothing.

DelSite is the third biotech firm to go bankrupt this year, according to figures from biobusiness magazine *BioCentury*. If that pace continues, 2009 will surpass the eight companies that went belly up last year (*Nat. Biotechnol.* **27**, 3–5, 2009).

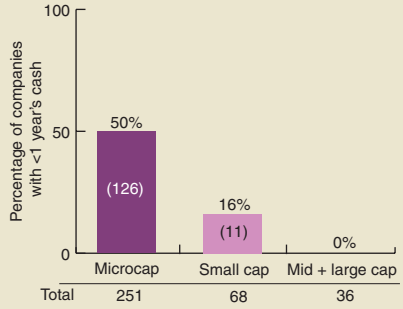


Figure 1 Percentage of biotech firms operating with less than one year's cash, segmented by market cap. Microcap, <$250 million; small cap, $250 million to <$1 billion; midcap, $1 billion to <$5 billion; large cap, ≥$5 billion.

In fact, a host of indicators suggest this year is shaping up to be worse than 2008. Ernst & Young has tracked the 'cash runway' of biotechs for years and usually finds between 20% and 25% of public companies have less than a year's cash. But *Nature Biotechnology* examined the most recent earnings reports (fourth quarter 2008 and first quarter 2009) of 355 global public firms that most closely met our definition of a biotech company and found that ~39% of them have less than one year's worth of cash. The increased percentage is driven for the most part by the plight of microcap firms (**Fig. 1**).

The shrinking market caps and depressed stocks have kept the exchanges busy with delistings, particularly NASDAQ. Eleven biotech companies have been delisted for regulatory issues or noncompliance through the first four months of this year, meaning 2009 could see >30 companies removed from the exchange (22 were removed for these reasons in 2008.)

The lack of investment has firms dumping programs to save on R&D and cutting staff, too. Over the six-month period leading to the end of March, *BioCentury* data show some 30 firms closed R&D programs in non-core areas. Restructurings so far in 2009 far outstrip those of previous years (**Table 1**).

These numbers paint an unpleasant picture, but it is still hard to draw long-term conclusions about the health of small biotechs. On one hand, it's quite likely that investors' value perception of biotech has changed permanently. Yet, it's also true that the need for biotech's strongest offerings (innovation and healthcare products) has not diminished. Regardless, the vaults will not open tomorrow, or next week, or even the third quarter, so biotechs should prepare for a trip through the desert.

Table 1 Public companies restructuring

Year	Number of public companies announcing restructurings
2006	35
2007	57
2008	114
2009	52 (first quarter alone)

Source: *BioCentury*.

That isn't to suggest massive death. "We don't think all those firms with less than a year of cash will disappear; we actually think the industry is quite resilient," says Glen Giovannetti, the global biotech leader at Ernst & Young, noting the restructuring and pipeline reduction happening across the sector. Still, there will be firms who "run out of options and end up shutting doors," he says. In fact, he expects more liquidations and bankruptcies "this time around" than in previous low times.

Exactly how many isn't clear, but if nearly 40% of public biotechs are in a cash crunch, Giovannetti estimates 20–25% of these could go under in the next 12–18 months. If that's the case, the international biotech sector could lose 25–35 more firms. The longer the economy languishes where it is, the higher that number could rise.

Many argue that the downturn is in effect culling the weak—Darwinian principles applied to biotech—and Giovannetti agrees, saying a "stronger cohort of companies" will come out the other side. The question is, are we looking at simple Darwinian selection or a mass extinction of microcaps similar to a cataclysmic event?

Brady Huggett, Senior Editor

IN brief

Mixed news for Avastin

Roche

Avastin's trial failure could have pushed down Genentech's stock.

It came as a shock, on April 21, when Roche of Basel announced that a highly anticipated phase 3 trial of Avastin in early-stage colon cancer had missed its primary endpoint. The humanized monoclonal antibody (mAb; bevacizumab), developed by the S. San Francisco, California–based Genentech, is a vascular endothelial growth factor (VEGF) blocker, and the company's best selling product. A successful trial for early-stage colon cancer would have boosted drug sales considerably, but results from the C-08 trial evaluating Avastin in combination with chemotherapy after tumor resection in 2,700 patients showed that the mAb failed to reduce the risk of recurrence. This proved a major disappointment to the Swiss pharma, which had only a month earlier completed a $46.8 billion takeover of Genentech. News of the trial failure sent Roche's shares tumbling and instigated talk that the Swiss company may have paid too much for the biotech. Had the results been known at the time of closing the deal Roche might have bought Genentech at a lower price. The good news came on May 5, when the US Food and Drug Administration approved Avastin as a therapy for recurrent glioblastoma multiforme in patients with refractory progressive disease. Avastin's approval for glioblastoma, an indication worth ~$300–$400 million per year, according to New York City–based senior biotech analyst George Farmer of Canaccord Adams in Vancouver, British Columbia, won't offset its loss in early-stage colorectal cancer for which analysts had estimated $1 billion per year in additional revenue. That Avastin is not active as an adjuvant in early colorectal cancer is perplexing, considering the success the drug has enjoyed in treating late-stage metastatic colorectal cancer, as well as advanced lung and breast cancers. "We know very little about the role of VEGF in the early stages of cancer progression," says cancer biologist and translational investigator Rakesh Jain, director of the Edwin L. Steele Laboratory for Tumor Biology at Harvard Medical School in Boston. "VEGF is just the first line in making blood vessels in tumors, and there are other pathways and growth factors including inflammatory cytokines that are needed for making blood vessels and that contribute to tumor progression and metastasis." Despite the setback, Roche is committed to testing Avastin in other programs, and in early-stage cancer. "But the potential of Avastin in other adjuvant settings, including breast and lung cancers, is questionable as well," says Farmer. "Now in hindsight it looks like Roche overpaid, based on the outcome of that [C-08] study."

George S Mack

Ariad's NFκB patent claims shot down on appeal

On April 3, a federal appeals court ruled in favor of Eli Lilly in Indianapolis, Indiana, capping a seven-year legal fight with Ariad Pharmaceuticals over its patent claims surrounding nuclear factor kappa B (NFκB). The case has embroiled big pharma, several prominent biotech companies and even Nobel Prize winners as courtroom witnesses in a high-stakes tussle over the commercial reach of patents with especially broad claims. The recent verdict reverses a 2006 jury ruling in favor of the Cambridge, Massachusetts–based Ariad, a decision that had, back then, surprised and alarmed many in the biotech industry (*Nat. Biotechnol.* **24**, 737, 2006). The dispute centered on Lilly's osteoporosis drug Evista (raloxifene) and sepsis drug Xigris (activated protein C), but the case's implications extended far beyond these two drugs. "Ariad sought to assert claims that are broad far beyond the scope of the [patent] disclosure," wrote appeals court Judge Kimberly Ann Moore in her opinion reversing the jury verdict.

The patent's claims were indeed broad. They encompassed all methods for lowering cellular levels of NFκB, a transcription factor involved in inflammation. NFκB was discovered in 1986 by Nobel Prize winner David Baltimore, then at the Massachusetts Institute of Technology (MIT). MIT, Harvard University and the Whitehead Institute, all of Cambridge, Massachusetts, together shared the patent. Because NFκB is so important in biology and disease—it has been implicated in arthritis, cancer, diabetes and stroke—the claims essentially gave Ariad, who had gained an exclusive license for the patent from the Cambridge group in 1991, patent rights over scores of marketed and experimental drugs that acted, directly or indirectly, on the NFκB pathway.

Ariad used the license aggressively, suing Lilly for infringement the day the patent issued in 2002, and sending letters to about 50 other companies asking them to license the patent. Lilly fought back hard, as did Amgen in Thousand Oaks, California, which filed a preemptive suit against Ariad in 2006 to invalidate the patent and certify that its rheumatoid arthritis drug Enbrel (etanercept) does not infringe.

The Amgen case is still unresolved, but Lilly's victory appears decisive. Despite the earlier jury verdict in Ariad's favor, "the federal circuit [court] treated these claims, you know, almost derisively. They just smacked them," says Minnesota patent attorney Warren Woessner, former chair of the biotech committee of the American Intellectual Property Law Association. Woessner had predicted Ariad's defeat. "They won in a jury trial—big deal. They got some Nobel prizewinners up there to say how wonderful this was, and the jury folded like a cheap lawn chair. That's not uncommon. But the [appeals judges] just demolished this."

In Woessner's mind, Ariad was unlikely to prevail in the long run, given past decisions of the US Court of Appeals for the Federal Circuit (known simply as 'the federal circuit'). Allowing such broad claims "was essentially impossible under federal circuit precedent," he notes, adding, "It just wasn't going to happen." Woessner

Lilly headquarters—the company has won a legal dispute with Ariad over patent rights surrounding NFκB.

says accepting Ariad's claims would have been like accepting "a claim on antigravity." In 2003 the federal circuit faced a similar case—the *University of Rochester v. Searle*. The university, which obtained a cyclooxygenase type 2 (Cox-2) patent but did not describe specific Cox-2 inhibitors, sued Skokie, Illinois–based Searle (now part of Pfizer) over Searle's cyclooxygenase type 2 inhibitor Celebrex (celecoxib). The university lost.

Ariad and its university co-plaintiffs took the same road, with the same outcome. Patent claims, to be allowable, must be supported by a written description of the invention detailed enough "to enable any person skilled in the art" to make and use the invention—a key requirement of US patent law. Although Ariad claims its patent, unlike the Rochester patent, discloses "specific information and specific guidance," the patent failed the written description test, and the court didn't even bother ruling on the enablement requirement.

The inventors "didn't do anything to enable even an iota of this particular patent," says Arti Rai, a law professor and patent expert at Duke University in Durham, North Carolina. Although the discovery of NFκB was a significant achievement, Rai says, it didn't give the discoverers the ability to lay claim to all future modulators of that pathway.

Ariad isn't admitting defeat as yet. The company's CEO Harvey Berger, in a press release, noted that the April ruling only invalidated four patent claims (out of 211) and invoked "only one of the technical requirements for validity." "We believe that this decision may allow us to pursue further legal action and review of the ruling," Berger commented in the release.

But Ariad looks beaten. Woessner predicts that "they're not going to get any further judicial review." In addition, the company's dispute with Amgen is in trouble, with Ariad appealing a September 2008 district court ruling that cleared Amgen's Enbrel of infringement. What's more, an ongoing US Patent and Trademark Office reexamination of the patent gives scant hope, as 157 of the 211 patent claims had been either rejected or cancelled as of March 16. Ariad could again sue for infringement based on the surviving claims, but it would face the same legal objections that proved fatal in the Lilly case.

The investment community isn't counting on any future royalties. "We expected the [original] ruling to be overturned," says Phil Nadeau, a biotech analyst for Cowen and Company in New York. "There was no value in Ariad stock for any royalty payments they could have received based on these patents."

So for the moment, broad upstream "mechanism of action" patents, like the NFκB patent, do not seem to pose much of a threat. "The pharmaceutical industry has had a bad run as plaintiffs in patent infringement cases, but they've been doing okay as defendants," says Rebecca Eisenberg, a law professor and biotech patent expert at the University of Michigan in Ann Arbor. "The federal circuit has been with them on invalidating these upstream patents that they've been charged with infringing. And the Supreme Court also."

Companies asserting broad claims "are not going to get much sympathy" from the federal circuit, agrees Rai. "And if they're trying to assert them against a defendant who is as willing to fight as Eli Lilly is, they're ultimately going to lose."

Many universities, however, emboldened by Ariad's 2006 district court victory, have been pressing for such broad claims. "Every professor that discovers a mechanism of action now wants you to claim it," says Woessner, who advises universities. "And it can be hard to dissuade them from that." The take-home lesson from the Ariad case, says Woessner, is that filing such broad claims, without specifying compounds, hoping that some will stand, is a risky patent strategy. "Don't try to get broad functional claims, like the Ariad claims, or the Rochester claims," he says, without describing specific pathway modulators.

There's a broader lesson in the NFκB dispute. In Rai's view, the case highlights the potential harm that universities can inflict when their patents broadly claim downstream commercial products. She points out that the 1980 Bayh-Dole Act, which granted universities ownership of patent rights, was intended to promote commercialization of federally sponsored inventions, not to place a tax on innovation by others. But Ariad, a reputable science-based biotech company, never tried to develop NFκB inhibitors on its own. Instead, it sought to use the license to collect a revenue stream from other companies. If such claims were allowed to stand, they could ultimately chill product development because companies developing novel drugs would face possible infringement from the outset—not a conducive mindset for undertaking risky drug development.

The universities holding the NFκB patent, in Rai's view, are ultimately at fault for the misuse of its license. (MIT made the licensing decision, but declined to comment for this story.) The NFκB patent "shouldn't have been applied for with that breadth," Rai says, "and then it shouldn't have been exclusively licensed, given that it was so broad, to one company that didn't seem to have the capacity really to develop it." In the end, MIT, Harvard and the Whitehead may receive very little from what remains of the contentious patent.

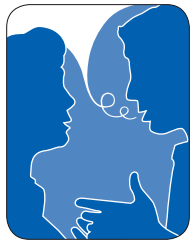
Ken Garber *Ann Arbor, Michigan*

IN brief

TNF-blocker triple approval

A new tumor necrosis factor alpha (TNF-α) blocker with a unique once-monthly dosing schedule has been approved, but despite its advantages, few believe it will shake up the market. Simponi (golimumab) won approval from the US Food and Drug Administration for three rheumatology indications—rheumatoid arthritis, psoriatic arthritis and ankylosing spondylitis—in April, and from Health Canada earlier that month. Simponi, a fully human anti-TNF-α monoclonal antibody produced by Johnson & Johnson's (J&J) subsidiary Centocor Ortho Biotech of Horsham, Pennsylvania, and Schering-Plough of Kenilworth, New Jersey, must compete in the already crowded rheumatology space, which includes J&J's own blockbuster Remicade (infliximab). Market watchers, however, believe it is unlikely Simponi will displace best-selling counterparts Enbrel (etanercept), Humira (adalimumab) and Remicade. Janice M. Reichert, a senior research fellow at Tufts Center for the Study of Drug Development in Boston, who collects data on emerging drugs in the industry says: "Remicade has an established market and it is difficult to push something out of [that] position." If a patient is responding well to conventional treatment, Reichert notes, the physician will be reluctant to switch to a new therapy, especially when a clear competitive advantage is lacking. Simponi's once-monthly dosing schedule, less frequent than that of other TNF-α blockers. could provide that advantage. *James Netterwald*

IN their words



"You need to live with that executive team. You need to be with that team."

San Francisco-based Corey Goodman insinuates the motives behind his resignation as leader of Pfizer's Biotherapeutics and Bioinnovation Centre, as the recently merged Pfizer-Wyeth executive teams locate to the East Coast. (*San Francisco Business Times*, April 29, 2009)

"It's fair to say that at some point the virus passed through a pig. It could have been months; it could have been years ago."

Paul A. Offit, an infectious disease expert at Children's Hospital of Philadelphia explains that, based on the virus's genetic structure, the animals do not seem to be playing a role now. (*New York Times*, April 28, 2009)

"We shot ourselves in the foot."

Paul Collier, professor of economics at the University of Oxford, on how a decade ago Europe, followed by Africa, banned GM crops, which now seem to offer a way to adapt to global warming. (*The Independent*, April 18, 2009)

IN brief

Phase zero launch

Scientists at the US National Cancer Institute (NCI) in Bethesda, Maryland, have conducted the first phase 0 oncology trial, which they claim could help accelerate drug development. The NCI phase 0 study tested a single dose of Abbott's small-molecule candidate ABT-888 in 13 patients with advanced cancers. The results, obtained in five months, proved that the drug from the Abbott Park, Illinois, company is well tolerated and inhibits its target—the poly(ADP-ribose) polymerase (PARP) enzyme—in tumor samples and blood cells. In a phase 0 study, a small number of patients is treated with nontoxic microdoses or a single dose of a new drug to obtain pharmacodynamic and pharmacokinetic data. Phase 0 trials can help identify and discard inactive drugs early in the process, ultimately improving success rates. But they are unlikely to become routine in cancer drug development, says Susan Galbraith, vice president oncology discovery medicine and clinical biomarkers at Bristol-Myers Squibb. "There are other approaches that can achieve the same goals faster and for less cost." Nicola Curtin, professor of experimental therapies at Britain's Northern Institute for Cancer Research (Newcastle), says a phase 0 study can be rolled in with a phase 1 to obtain the same data. For monoclonal antibodies, microdosing may present a challenge. "Traditional antibodies have a 1–2 week half-life so one dose of those drugs would give equivalent exposure to multiple doses of a small-molecule drug," Galbraith points out. *Emma Dorey*

Pig patent revolt

Germany, home to the biggest swine population in Europe, is up in arms over a patent covering a marker-assisted test to breed meatier pigs. The patent covers a screening method to identify a polymorphism in the leptin receptor gene, useful for selecting animals for stockbreeding. The patent, originally filed with the EPO by Monsanto, was granted last July to Newsham Choice Genetics, the West Des Moines, Iowa–based company that in 2007 acquired Monsanto's porcine genetics subsidiary. The gene sequences and the test kit itself, although originally included in the application, were not part of the patent granted by the European Patent Office (EPO). Several notices of opposition have been filed, mainly from nongovernmental organizations and individuals, not by competing companies. On April 15—the day before the deadline for objections—activists and farmers demonstrated outside the EPO's Munich office, protesting about paying royalties to a US firm. "This seems like a complaint from the 18th century," says Larry Schook, co-chairman of the Swine Genome Sequencing Consortium, which will be completing its sequencing effort by August. According to Schook, breeding companies often sell germplasm with dubious genetic merit at a premium. Marker-assisted tests will offer an actual genetic benefit rather than a proposed one. Gordon Wright, from the Chartered Institute of Patent Attorneys in London, speculates the company "will be aiming to enforce the patent against commercial [kit] suppliers rather than breeders." *Anna Meldolesi*

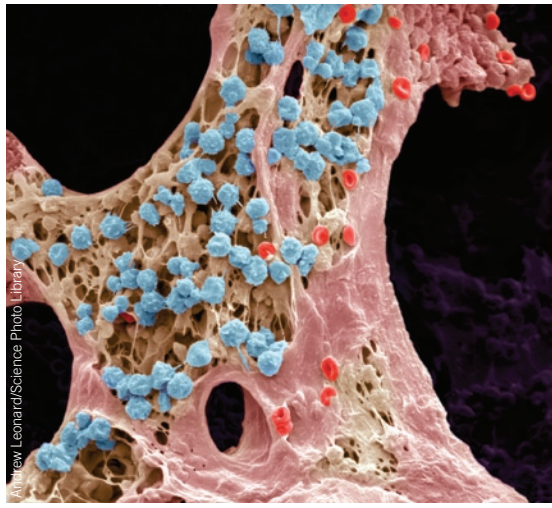
US regulator wades into stem cell therapies for heart disease

The US Food and Drug Administration's (FDA) draft guidance on cell therapy for cardiac disease has been given a muted welcome by clinicians and industry—not least because it may bolster the reputation of a field that thus far has enjoyed more than its fair share of charlatans and quacks. One impetus for the April release of the guidelines is problems cropping up with existing bone marrow–-derived cell therapies for heart disease in the clinic. "My intuitive sense is that they'd had some issues [with companies doing cell therapy heart research] already," says attorney Edward J. Allera, chairman of Buchanan Ingersoll & Rooney's food and drug group in Washington, DC. "Educated players understood this was coming," he says.

Elmar R. Burchardt, vice president of medical affairs for Aastrom Biosciences, a regenerative medicine company based in Ann Arbor, Michigan, has a different take. "It could be a sign that the field is maturing—that the FDA thinks this is an important emerging field," he says.

Clinical work with cell therapy in heart disease has been under way in Europe since 2001, although "in [the US], it took a little longer," Burchardt points out. In that year, a pair of seminal articles (*Nature* **5**, 701–705, 2001 and *Nat. Med.* **7**, 430–436, 2001) established that bone marrow–derived cells could be used to repair damage after heart attacks. Data from the European studies surfaced in 2003, and much work afterwards went into defining the target populations. Should cells be used alone or as an add-on procedure with bypass grafting? At which stage of the disease should transplants be considered? When and how should the bone marrow–derived cells be injected? Are selected cell types better than unselected bone marrow? Meanwhile, the FDA "had a lot of hearings," grappling with uncertainty over matters as basic as whether trials should be controlled or open label.

The guidance laid out by the agency's Center for Biologics Evaluation and Research is aimed at steering preclinical and clinical studies (**Table 1**), and states the information needed to back up investigational new drug applications. "It's a very different game [now]," Burchardt says. The game's rules to be, or at least a step toward them, are outlined in the nonbinding recommendations drawn up by the agency, and concerned parties have until July 1 to submit their remarks. Aastrom is not tipping its hand regarding how the company will respond, though Dan Wolin, the company's manager of regulatory affairs,



Bone marrow is a rich source of adult stem cells. A number of companies are delivering such cells for cardiac repair.

SELECTED research collaborations

Partner 1	Partner 2	$ (millions)
Kyowa Hakko Kirin (Tokyo)	Sanofi-Aventis (Paris)	315
Isis (Carlsbad, California)	Alnylam (Cambridge, Massachusetts)	31
Institute of Ophthalmology, University College (UCL) London	Pfizer Regenerative Medicine (Cambridge, UK)	*

*Not disclosed

Table 1 Selected companies with cardiac cell therapies under development

Company	Drug	Treatment details	Stage of development	Indication
Osiris (Columbia, Maryland) Genzyme (Cambridge, Massachusetts)	Prochymal	Adult mesenchymal stem cells obtained from the bone marrow of volunteer donors formulated for intravenous injection.	Phase 2	Coronary artery disease
Baxter (Deerfield, Illinois)	Autologous CD34$^+$ stem cells	Autologous CD34$^+$ stem cells gathered from the patient's blood. CD34$^+$ cells are enriched by separation and injected using a catheter-based, nonsurgical system into heart areas with poor blood flow.	Phase 2	Refractory chronic myocardial ischemia
Aastrom (Ann Arbor, Michigan)	Tissue repair cells	An adult mixed-cell product containing stem cells and progenitor cells from a patient's own bone marrow. The autologous cells are prepared using the single-pass perfusion system and expanded over 12 days before transplantation.	Phase 2	Congestive heart failure
T2cure (Frankfurt, Germany)	t2c001	Autologous bone marrow–derived progenitor cells, infused into the desired vessel through a catheter.	Phase 1/2	Congestive heart failure
Athersys (Cleveland, Ohio)	MultiStem	A bone marrow–derived off-the-shelf stem cell product	Phase 1	Acute myocardial infarction

predicts that the dual device–drug aspect of the research one of the main concerns for researchers in general, along with the problem of how to decide on trial endpoints.

At present, most autologous cellular therapies for cardiac disease work in the following manner: a small aspirate is collected from a person's bone marrow, the harvested cells are separated and enriched for CD34$^+$ stem cells, and reinjected into the patient's heart. The mechanism by which these cells exert their benefit remains controversial, but most agree it is down to growth factors that prevent the cell death that typically follows a myocardial infarction, and, at the same time, encourage revascularization. The choice of endpoints for each study should reflect the expected mechanism of action and the indication, the guidelines state. Most studies assess left ventricle ejection fraction, which reflects the blood pumped out with each heartbeat, a common measurement of overall heart function that typically declines after a heart attack, but in general, trial participants should be followed for survival.

How to deliver the cells is another consideration. Brian Bruckner, principal investigator in the IMPACT trial conducted by Aastrom at the Methodist DeBakey Heart & Vascular Center in Houston, injects patient-specific cells through the pericardium, and into 25 areas around the left ventricle. Osiris Therapeutics, of Columbia, Maryland, delivers its Prochymal, a formulation of mesenchymal stem cells, through a standard intravenous line. In early April, Osiris treated its first heart-attack patient in a phase 2 trial. Deerfield, Illinois–based Baxter uses an endocardial approach—a catheter inside the heart. Six-month data from a phase 2 trial proved positive in late March.

"There are different mindsets on how to deliver cells, and [users of the catheter approach] have come up with some proposed mapping systems based on [magnetic resonance imaging] and electrical conductance," Bruckner says. Catheter-based delivery has the advantage of avoiding general anesthetic in an operating room. "But with a catheter, you don't have a lot of control on where they [the cells] go. You're inside a beating heart with blood moving everywhere. Some [cells] could be washed out by the turbulent blood flow. You don't know if the cells are getting where you want them to be, and are going to stay." It's these problems—along with the ever-tricky chemistry, manufacturing and controls aspect of devising somatic cell therapies—that the draft guidance is designed to address.

At first, the FDA seems to have been satisfied with the groundwork laid down by European researchers, allowing Aastrom to forgo phase 1 trials in the US. The downside is that cardiac cell therapy's technology has been "rushed," notes Bruckner. "I look at stem-cell therapy as skipping the basic science that needed to be done before it was brought to the public."

Issuance of the draft guidance, though, suggests the FDA is tightening its scrutiny. Cell therapy experiments in cardiac disease have also been conducted "offshore, and that's raised some concern," attorney Allera says. "People are looking at doing this globally and bringing the data back. But there are rules here. You can't just jump in," though the document does not explain in much detail what the rules for cardiac cell therapies are. Allera adds that regulators are taking the customary route of making sure "they have enough flexibility [in the issued guidance] that they can turn you down."

Susan Cruzan, spokesperson for the agency, admits that the regulatory route for cell-therapy products (e.g., biologic, device) is "a complex one, and one size will not fit all." She

Details
Kyowa Hakko Kirin has granted Sanofi-Aventis Group exclusive, worldwide rights to develop and commercialize a human monoclonal antibody (mAb) against tumor necrosis factor ligand superfamily member 14 (TNFSF14, LIGHT, CD258), except in Asia where the companies will co-develop the product. The anti-LIGHT mAb is in preclinical development to treat ulcerative colitis and Crohn's disease. Kyowa, which retains marketing rights, could receive up to $315 million in up-front and milestone payments, plus royalties.
Alnylam is paying Isis up to $11 million up-front to license its single-stranded (ss)RNA interference (RNAi) technology. Under the terms of the agreement, Isis will receive up to $10 million on demonstrating *in vivo* efficacy in rodents, $5 million after demonstration of efficacy in nonhuman primates and $5 million at the start of the first clinical trial with an ssRNAi product. Both companies will be allowed to develop drugs with the new technology.
Pfizer and UCL have entered a collaboration and license agreement to develop stem cell treatments for wet and dry macular degeneration as well as other retinal diseases. Under the terms of the agreement, Pfizer will provide funding to UCL researchers to understand how human embryonic stem cells can differentiate into retinal pigment epithelium (RPE). Pfizer gains exclusive worldwide rights to develop and commercialize any resulting product for use in the ophthalmology field.

IN brief

Genzyme takes Campath bet

Genzyme has shored up its oncology and multiple sclerosis franchise through a new deal with Bayer, of Leverkusen, Germany. Best known for targeting rare genetic disorders, Genzyme is bringing into its stable three approved cancer therapeutics: Fludara (fludarabine) for B-cell chronic lymphocytic leukemia (B-CLL); Leukine (sargramostim), a hematopoietic growth factor used to stimulate the bone marrow following chemotherapy; and Campath (alemtuzumab) for B-CLL. The Cambridge, Massachusetts, company will pay up to $500 million (plus another $150 million after 2011) based on annual revenues, for those three products. Bayer will continue to supply Fludara and Leukine, although Genzyme is acquiring a Leukine-manufacturing plant for $75 million to $100 million and will produce that product itself when the plant is cleared by the US Food and Drug Administration. Perhaps the most interesting aspect of the deal is that Campath is being tested against multiple sclerosis (MS). The companies have been linked over Campath since 2006, but this deal transfers all marketing rights to Genzyme while Bayer continues to support development. Campath, a humanized monoclonal antibody that binds to CD52, has shown promise in phase 2 trials for MS, and is currently in two phase 3 trials. If approved for MS, Bayer could receive up to $1.25 billion, plus further payments related to sales after 2021 (Genzyme retains a buyout option for Bayer's share in 2020 for $900 million). Although Bayer could still co-promote the drug in the US, Genzyme now has primary responsibility in MS. But don't let the high milestone payments deceive—all those moving parts in essence mean the deal is more about "altering accounting and delineating development responsibilities" rather than making a strategic shift for Genzyme, says Brian Abrahams, analyst with Oppenheimer in New York. *Brady Huggett*

New product approvals

Removab (catumaxomab)/Trion Pharma (Munich)/Fresenius (Homburg, Germany)

The European Commission on April 23 approved Removab, a new-generation trifunctional antibody, to treat malignant ascites common in ovarian, pancreatic and gastric cancers. The antibody possesses two different antigen binding sites: one targets the human epithelial cell adhesion molecule (EpCAM), which is found in the majority of epithelial tumors, the other targets human CD3 on T-lymphocytes.

Simponi (golimumab)/Johnson & Johnson (New Brunswick, New Jersey)

The US Food and Drug Administration on April 24 approved Simponi, a second-generation tumor necrosis factor-α inhibitor, to treat three forms of arthritis. Simponi is a once-monthly subcutaneous injection for treating adults with moderate to severe rheumatoid arthritis, active psoriatic arthritis and active ankylosing spondylitis.

points to a clause in the guidance that recommends sponsors "consult with FDA concerning the regulatory pathway for the use of cell selection devices." Regarding any developments that might have prompted the guidance, Cruzan says that the agency "wouldn't be able to comment on the safety risks observed in trials performed thus far, due to confidentiality regulations or anything under an [investigational new drug application]."

In all cardiac cell therapy, maintaining purity, potency, identity, sterility and shelf life—all elements of cell therapeutics as called for in the guidance—is the key to a regulatory win, along with proving what the mechanism of action is, said Andrew L. Pecora, chairman of Amorcyte's board, a company based in Hackensack, New Jersey, that has recently released phase 1 results of an autologous stem cell treatment for heart muscle damaged by infarcts. Pecora claims the overall task is not as difficult as it seems. "If I'm going to squirt cells through a three-foot catheter into your heart, you want to know they're going to be okay when they come out the other end. There are some very simple things [to determine], that is, is your project sterile? Is it pure? Some areas are going to be fairly easy," whereas others will be more challenging, such as batch-release specifications for each autologous group of cells. "Instead of having one huge batch in a vat, with cell therapy, it's going to be real-time delivery. There's an infrastructure that's different from [work typically done by] the biotech or pharmaceutical industry."

One thing seems clear: the agency intends to classify cell therapies as combination products. The FDA's Center for Biologics Evalution and Research "would be the lead division and consult with the device group," called the Center for Devices and Radiological Health (also FDA). "In the small-molecule world, this happens as well," says Wolin—and it adds an extra hurdle. More daunting, perhaps, are the requirements for cell therapy manufacture. "Several animal studies and/or species may be necessary to adequately model functional aspects and potential toxicities of a single product," the draft says, but how many will be required will depend on the "biological characteristics of the product," a point likely to be argued over.

Each cell-therapy company will have its own conversation with regulators, says Paul J. Schmitt, acting CEO of Amorcyte. Schmitt lauded the draft guidance, saying the field "needs to progress from outstanding science to the rigors of real drug development," a move that would "start to take the charlatans out of the industry."

Randy Osborne *Mill Valley, California*

India's first true stem cell trials

The Drug-Controller General of India (DCGI) has given the go-ahead for the first clinical trials designed to test stem cell products. Stempeutics Research of Bangalore launched a combined phase 1 and phase 2 trial on April 22 to evaluate whether its stem cell products can benefit people who have experienced myocardial infarction and individuals with critical limb ischemia (CLI)—a condition that often requires amputation.

"These are the only two stem cell trials officially approved to date," says Polani B. Seshagiri, a member of a government panel that made the recommendation to DCGI. There have been many claims in the past from Indian labs offering stem-cell therapy to treat a wide range of diseases, prompting criticisms that local regulators were failing to monitor the procedures (*Nature* **434**, 259, 2005). But Seshagiri, who heads the Stem Cell and Transgenic Research Lab at the Indian Institute of Science in Bangalore, says, "none of these can be called a clinical trial." The Stempeutics trials are the first randomized, double-blind, multicentric, placebo-controlled studies. Each trial will recruit 100 patients in batches.

"Our goal is to bring out affordable stem cell–based products as drugs in chemists' shops," says Stempeutics' president Balu N. Manohar. The company extracts mesenchymal stem (MS) cells from the bone marrow of healthy donors and expands them in culture before infusing them back in. These MS cells are well tolerated by the recipient because they lack immunogenic major histocompatibility complex class II molecules on their surface, says Ramesh Bhonde, technical director for Stempeutics. The company has scaled up the production process to obtain 300 to 400 million MS cells of good manufacturing practice quality from a single donor. Preclinical animal toxicity studies, says Bhonde, confirm that *ex vivo* cultured adult MS cells are safe and can be used both in autologous and allogenic settings. Seshagiri does not expect any major obstacles in this form of therapy as long as the cells are clinical grade.

Killugudi Jayaraman, Bangalore, India

Vaccine market boosters

Recent commercial success belies conventional wisdom that vaccines are a low-margin, moribund sector. But will the trend continue? Cormac Sheridan investigates.

GlaxoSmithKline (GSK) Biologicals reached an important milestone on March 30 in its effort to make up lost ground on rival vaccine maker Merck. Almost two years to the day after submitting a biologics license application (BLA) to the US Food and Drug Administration (FDA), London-based GSK filed final phase 3 study data for its cervical cancer vaccine Cervarix (recombinant adjuvanted bivalent human papillomavirus (HPV) vaccine). Although approved in Europe since September 2007, blockbuster status has eluded Cervarix so far. And it is by no means clear that it will be able to gain significant market share in the US, given the blitzkrieg marketing tactics adopted by Merck, of Whitehouse Station, New Jersey, in rolling out its vaccine Gardasil (recombinant quadrivalent HPV vaccine).

A market with historically low margins, the vaccines business has recently been reshaped by a handful of highly priced blockbuster products (**Table 1**). The approvals of Prevnar (pneumococcal septavalent conjugate vaccine) in 2000 and of Gardasil six years later have shaken up this previously underperforming segment of the pharmaceutical market. And although the blockbuster model in the wider pharmaceutical industry may be in decline, blockbusters are gaining prominence in the vaccine sector. "Contrary to popular opinion, the blockbuster model still exists in vaccines," says John Savopoulos, director of VacZine Analytics, a specialist vaccines market research firm based in Bishop's Stortford, UK. With vaccine sales slowing dramatically in 2008, the question is, Are there other vaccines in the pipeline to spur another wave of double-digit growth?

Getting approval for pediatric use can be key to achieving blockbuster status for some new vaccines. Source iStockphoto.com

Winners and losers

Much of the 2008 slowdown in sales can be attributed to Gardasil, which appears to have saturated its home market already, after having gained a unanimous recommendation from the US Centers for Disease Control's advisory committee on immunization practices following its approval in June 2006 by the FDA. Merck reported $1.4 billion in sales for the vaccine in 2008, versus $1.5 billion in 2007, its first full year on the market. It is forecasting US sales of $1.1 billion to $1.3 billion in the current year, although first quarter sales of $262 million disappointed analysts. "Gardasil is beginning to slow down now. It's matured quite quickly," says Savopoulos. (These figures omit European sales through Sanofi Pasteur MSD, a fifty-fifty joint venture between Merck and Sanofi Pasteur, the vaccines arm of Paris-based Sanofi Aventis. That entity added another €584 ($776) million in sales of the product in 2008.)

Slower growth of seasonal flu vaccines also affected overall growth figures last year, as did the weak US dollar, according to VacZine Analytics. Leading European firms, such as Novartis of Basel, GSK and Sanofi Pasteur, report sales in their respective national currencies, so currency conversion effects added an unflattering gloss to their US sales figures. The vaccines industry, which is driven by large economies of scale and deep manufacturing expertise, is strongly oligopolistic. Rounding out the top five players is Prevnar's marketer, Wyeth, of Madison, New Jersey. Between them, they command around 85% of the total market and had combined sales of $17.5 billion in 2008, according to VacZine Analytics. Sanofi Pasteur MSD, which reports separately from both parent companies, added another €1.3 billion. The overall total represents an improvement of little more than 5% over 2007, in stark contrast to the massive 40% hike in sales achieved during 2007, according to the same source, largely because of Gardasil.

If Merck and Wyeth (soon to become part of New York–based Pfizer) have taken the vaccine industry's biggest prizes during the past decade, GSK has been the biggest loser. Long touted as another blockbuster, Cervarix attained modest sales of £125 ($177) million in 2008, its first full year on the market. Several factors could cause snags in the road to regulatory approval. The vaccine includes AS04, an adjuvant comprising the bacterial cell wall constituent monophosphoryl lipid A. Until now, "no novel adjuvant [has been] approved for use in the US," says Savopoulos, and approval of Cervarix would signal a thawing of attitude on the part of the FDA toward alternatives to the traditional adjuvant alum. What's more, Cervarix fails to offer protection against HPV strains 6 and 11, which cause genital warts. On the plus side, if GSK can show that Cervarix offers a more durable response than Gardasil, it could make substantial inroads into the latter's market share.

A new wave of blockbusters?

As the vaccines market continues to expand, the introduction of individual products will have a proportionately smaller effect on its growth. Nevertheless, the arrival of the next wave of blockbuster products is expected to deliver another big boost to the sector. Until then, the market will continue to grow at around 5–10% annually. Savopoulos identifies three potential growth markets: new pediatric vaccines, new travel vaccines and vaccines to protect against hospital-acquired infections, which would represent a completely new treatment paradigm.

It may be too early to predict who will dominate these newly emerging product segments during the next decade, but what is clear is that each will be more hotly contested than has previously been the case, as pharmaceutical companies commit more resources to what has become one of the industry's biggest growth markets. "It used to be a one-horse race," Savopoulos says. Each vaccine maker was previously able to lay claim to a given product segment. "What's happening now is it's becoming a two- or three-horse race, where everyone has to share." Moreover, biotech companies are also looking for a piece of the action, as vaccines—traditionally rather low-tech commodities—are hauled into the molecular biology era (**Box 1**). Biotechs also ride the wave when the fear of a pandemic rises, such as the outbreak of H1N1 in March. (see p. 493)

One or two contenders have already hit the front in some of the more important contests that are now taking shape (**Table 2**). In a recent analysis, Decision Resources, a Waltham, Massachusetts–based market research firm, identified MenB, a vaccine in development by Novartis Vaccines for preventing *Neisseria*

Table 1 Blockbuster vaccines

Company	Product	Sales revenues 2008
Wyeth	Prevnar	$2.7 billion (€2.02 billion)
Merck	Gardasil	$1.4 billion (€1.01 billion)
	ProQuad/M-M-R II/Varivax	$1.3 billion (€974 million)
GlaxoSmithKline	Infanrix/Pediarix	$1.3 billion (€974 million)
	Hepatitis vaccines	$1.2 billion (€899 million)
Sanofi Pasteur	Polio/whooping cough/Hib vaccines	$1.1 billion (€823 million)
	Influenza vaccines	$1.1 billion (€823 million)

Currency conversions are as of May 8, 2009. Source: Research and Markets, Dublin

meningitidis serogroup B infection, as one of the most promising vaccines in the industry pipeline[1]. It is forecasting peak sales of $2 billion to $2.5 billion for the product. The vaccine is currently in a pivotal phase 3 trial, whose results are due later this year. "It's come a long way. So far it looks very good," says Rino Rappuoli, global head of vaccines research at Novartis, who began work on the project in the mid-1990s.

The pathogen remains a major cause of bacterial meningitis in infants and adolescents, particularly since effective vaccines against two other causes of the disease, *Streptococcus pneumoniae* and *Haemophilus influenzae*, have become available. It also causes septicemia. "Pediatricians are eagerly awaiting the arrival of MenB, and so we believe the receptivity is going to be very high for this vaccine," says Hemali Patel, infectious diseases analyst at Decision Resources. Quadrivalent polysaccharide-based vaccines covering the other four pathogenic meningococcal serogroups, A, C, Y and W-135, have been available for adult use for several decades, and conjugated vaccines were introduced several years ago for adolescents. However, such an approach is not feasible for serogroup B, because it produces a capsular polysaccharide identical to a polysialic acid moiety found on many human glycoproteins.

Rappuoli and his team mined the bacterium's genome sequence to identify novel antigens capable of inducing a strong antibody response[2]. That work, he says, created more knowledge about the pathogen's antigenic makeup than the sum total of what had been known previously. "It's become a tool that was totally unknown ten years ago," says Rappuoli, who coined the term 'reverse vaccinology' to describe the approach. Five highly conserved antigens are included in its universal vaccine, which induced a bactericidal antibody response in mice in 78% of a panel of 85 meningococcal strains tested. That coverage, Novartis Vaccines announced, can be boosted to over 90% by adding such adjuvants as CpG oligonucleotides, which mimic bacterial DNA, or Novartis's proprietary MF59 adjuvant, an oil-in-water emulsion.

Key to the vaccine's commercial success, Savopoulos says, is obtaining a recommendation for infant immunization. "If they don't get that recommendation and are limited to the current recommendation, they won't get blockbuster status," he says.

Decision Resources also identified PCV13 (pneumococcal 13-valent conjugate vaccine), which Wyeth is developing, as another likely blockbuster. That product, the subject of a BLA filing on March 31, is a successor to Prevnar. It offers improved coverage against *Streptococcus pneumoniae*, although it by no means offers complete protection, as 90 distinct serotypes have been identified thus far. Even so, Wyeth already appears to have gained the upper hand in this particular contest. Earlier this year, GSK CEO Andrew Witty said his company would not attempt a US launch of its competing product Synflorix (pneumococcal decavalent conjugate vaccine), which gained a European approval on March 31. Unlike PCV13, Synflorix does not offer protection against *S. pneumoniae* serotype 19A, which has emerged as a major pathogen following the success of Prevnar in combating other common serotypes.

Cytomegalovirus (CMV) represents another major opportunity in the pediatric vaccines market. The virus, a member of the herpes virus family, affects an estimated 40,000 newborns in the US every year. It causes around 400 deaths and leaves about 8,000 children with permanent disabilities, such as hearing loss, blindness or mental retardation. A recent, investigator-sponsored phase 2 study of a vaccine comprising recombinant CMV envelope glycoprotein B and the Novartis-owned MF59 adjuvant reported 50% efficacy in terms of infection rates per 100 person-years[3]. That outcome was "a very pleasant surprise," says Robert Pass, professor of pediatrics at the University of Alabama, Birmingham, and lead investigator on the trial. The expectation was that the vaccine would alter—but not prevent—the course of the infection in women who contract the virus during pregnancy and thereby have a positive effect on the congenital infection. The study indicates that glycoprotein B is a "very important component of a CMV vaccine," he says. "What remains to be determined is whether glycoprotein B alone would be sufficient."

Box 1 Biotech investors embrace vaccines

Although Wyeth abandoned its reported $1 billion bid for Leiden, the Netherlands–based vaccine maker Crucell earlier this year—after Pfizer made a $68 billion bid for Wyeth—the episode did demonstrate that vaccine companies are in play. However, René Verhoef, analyst at Fortis Bank, in Amsterdam, says that Crucell itself is not likely to attract bids from other vaccine makers in the near term, because of its "relatively modest product pipeline" and its lack of progress in bringing any of its own products onto the market. Its marketed portfolio is based on its $448 million stock-based acquisition of Berna Biotech, of Bern, Switzerland, in 2005. "If you want to become a serious player in the market you have to bring products onto the market yourself. That's a step Crucell hasn't been able to make so far," says Verhoef.

At the other end of the investment cycle, meanwhile, three venture capital deals involving early-stage vaccine developers provide further evidence of investor interest. GlycoVaxyn, of Schlieren, Switzerland, raised CHF 25 ($21.8) million in a Series B round to advance its biology-based system for producing conjugate vaccines. Vivaldi Biosciences, of New York, raised $23 million in a Series A round to develop a live attenuated influenza vaccine for older adults. The approach is based on altering nonstructural protein 1, an important virulence factor of the virus. Genocea Biosciences also raised $23 million in a Series A round, to promote a platform for identifying pathogen-specific antigens that elicit protective T-cell responses. GlycoVaxyn is, like many of its big pharma counterparts, targeting *S. aureus*, among other pathogens. Its approach is based on using the cloned bacterial glycosylation apparatus (from *Campylobacter jejuni*) to produce conjugated vaccines in an *Escherichia coli* background. It is using known antigens, but avoiding the harsh chemical treatments currently used in the development of conjugated vaccines that can have an impact on their immunogenicity. "Everything is done biologically," says CEO Philippe Dro.

The success of Gardasil has created a commercial rationale for the development of other prophylactics, such as the CMV vaccine, targeting the pubescent section of the population. "There is a logical place in the market for a cytomegalovirus vaccine for young persons who are reaching sexual maturity," says Pass. It will take five or six years, however, before one will be ready for launch, he adds. Sanofi Pasteur gained ownership of the glycoprotein B–based vaccine from Chiron, of Emeryville, California, around 10 years ago and may now reformulate it with its own adjuvant before further trials. Novartis, which paid $5.1 billion to become outright owner of Chiron in 2006, entered—or re-entered—the space late last year by in-licensing an investigational vaccine CMV vaccine from AlphaVax, of Research Triangle Park, North Carolina. The deal involved an up-front payment of $20 million, plus a potential equity investment as well as undisclosed milestones and royalties. The AlphaVax candidate vaccine, which is due to enter a phase 2 trial this year, consists of an alphavirus replicon particle encoding three components, phosphoprotein 65, immediate early protein 1 and soluble glycoprotein B[4]. DNA vaccine specialist Vical, of San Diego, is also in this contest. It is running a phase 2 trial of its CMV vaccine, VCL-CB01, which encodes glycoprotein B and phosphoprotein 65 (ref. 5). GSK has a preclinical program underway as well.

Table 2 Selected vaccines in development

Company	Product	Target	Status
GlaxoSmithKline	Cervarix	HPV	Registration[a]
	Synflorix (decavalent conjugate vaccine)	*S. pneumoniae*	Approved[b]
Novartis	Menveo (pentavalent conjugate vaccine)	*N. meningitidis* serogroups A, C, W, Y	Phase 3
	MenB (five-component subunit vaccine)	*N. meningitidis* serogroup B	Phase 3
Wyeth	PCV 13 13-valent conjugate vaccine	*S. pneumoniae*	Registration
	RLP2086	*N. meningitidis* serogroup B	Phase 1 (infants) Phase 2 (adolescents)
Merck/Intercell	V710 subunit vaccine	*S. aureus*	Phase 2/3
Novartis/Intercell	IC43	*Pseudomonas aeruginosa*	Phase 2
Sanofi Pasteur	ACAM-CDIFF	*C. difficile*	Phase 2b
	Dengue vaccine (tetravalent vaccine)	Dengue virus	Phase 2b

[a]Under review at FDA; approved in over 90 countries. [b]Approved Canada, Australia and Europe.

Breaking new ground

Preventing hospital-acquired infections, involving pathogens, such as methicillin-resistant *Staphylococcus aureus* (MRSA) or *Clostridium difficile*, represents new ground for vaccine makers and could open up several lucrative market segments, given the levels of mortality and the high costs associated with the problem. The inexorable spread of antibiotic resistance has, obviously, created the opportunity for novel vaccines, as well as passive immunotherapies, in the form of monoclonal antibodies, but these will complement rather than replace antibiotics. "You will still need antibiotics going forward, but obviously vaccines could limit the use of antibiotics," says Gerd Zettlmeissl, CEO of Vienna-based biotech firm Intercell. For Savopoulos, the argument in favor of vaccines also stacks up. "It would be very easy I think to make a case to give someone a vaccine for a couple of hundred dollars," he says, rather than pay for a prolonged stay in the hospital with associated antibiotic treatment and other healthcare support.

It is clear, however, that such products will have to clear high safety and efficacy hurdles to justify the kind of premium pricing their makers are likely to seek, says John Lebbos, vice president, infectious diseases, at Decision Resources. One reason for this is "it's not really expensive to give somebody prophylactic antibiotics," he says, although this would raise other concerns such as the selection of drug-resistant strains.

Merck is in the lead on the *S. aureus* front at present, by virtue of a licensing deal with Intercell on V710. The vaccine is based on a single antigen, IsdB, a highly conserved, iron-sequestering protein that the bacterium expresses on its cell surface[6]. V710 is undergoing a phase 2/3 trial, which will recruit around 8,000 patients undergoing cardiothoracic surgery. Trials in other indications could follow. "The use could be even broader, we think, than just elective surgery," Zettlmeissl says. An interim analysis is expected this year, and the study is due to be completed in February 2011. A regulatory filing in that initial indication could follow later that year, if it's successful.

The challenge for any effective vaccine lies in the complexity of *S. aureus* infection. It is characterized by multiple virulence factors, which are differentially expressed during different phases of growth. Nabi Biopharmaceuticals, of Rockville, Maryland, had a high profile failure in a phase 3 trial in 2005 with a multi-component conjugate vaccine called StaphVAX, which was tested in 3,600 patients undergoing hemodialysis. A successor, PentaStaph (pentavalent *S. aureus* conjugate and toxoid vaccine), was the subject of a Cooperative Research and Development Agreement Nabi entered with the US military late last year.

Sanofi Pasteur has the industry's lead program in combating *C. difficile* infection, following its £285 ($425) million acquisition of Cambridge, UK–based Acambis in September 2008. Both the incidence and severity of *C. difficile*–associated disease, commonly connected with disruption of the normal gut flora by antibiotic usage, have increased sharply. "It's on more death certificates in the UK than MRSA," says Savopoulos. "Not many countries have a handle on it yet." Through the same transaction, Sanofi Pasteur also gained outright ownership of the industry's most advanced vaccine against mosquito-borne dengue fever, having previously been a partner on the program. The tetravalent vaccine, which combats all four *Flavivirus* serotypes that cause the infection, entered a phase 2 efficacy trial in children in Thailand earlier this year. The company has already built a production plant for the product. "To make a capex [capital expenditure] decision means they're confident," says Savopoulos. Unlike many other parts of the drug industry, confidence is a commodity that is widely available within the vaccines sector.

Cormac Sheridan, Dublin

1. Patel, H. & Makarova, A. *Emerging Vaccines* (Decision Resources Waltham, Massachusetts, December 2008).
2. Giuliani, M.M. *et al. Proc. Nat. Acad. Sci. USA* **103**, 10834–10839 (2006).
3. Pass, R.F. *et al. NEJM* **360**, 1191–1199 (2009).
4. Reap, E.A. *et al. Vaccine* **25**, 7441–7449 (2007).
5. Wloch, M.K. *et al. J. Infect. Dis.* **15**, 1634–1642 (2008).
6. Kuklin, N.A. *et al. Infect. Immun.* **74**, 2215–2223 (2006).

bioentrepreneur
frombenchtoboardroom

Can you hear me in the back?

Carin Canale

With thousands of biotech companies out there, and just as many corporate presentations, here's how to differentiate your enterprise in an increasingly crowded marketplace.

Whether the markets are bullish or bearish, the coffers full or empty, attracting financing is ever present at the top of the agenda for any emerging life science company. With competition on the rise and availability of private investment becoming more difficult to secure, it is critical that you stand out and present your story in a compelling yet succinct manner. There may not be a one-size-fits-all secret to obtaining venture capital, but there are certainly steps you can take to improve your odds. Clearly, certain nonnegotiable elements are needed to obtain venture capitalist (VC) interest, including valid technology addressing a significant, unmet need, a solid business plan with a realistic exit strategy and an experienced management team to carry out the plan. The challenge is to differentiate your firm from the thousands of life science companies who can also claim to meet those prerequisites—and the solution often lies in delivering an impressive investor presentation.

Just as there is no one-size-fits-all approach to obtaining venture capital, there isn't a single investor presentation that will work for both the 15-minute slots at investor conferences and the much longer one-on-one VC meetings. That said, the following 15-slide outline provides a foundation upon which investor presentations for all occasions can be built, along with a few basic principles of communication and advice from venture capitalists.

Table 1 The dos and don'ts of interacting with venture capitalists

Do	Don't
Research the venture capitalist (VC) in advance	Bash the competition
Pay attention to what you say during the presession banter	Hype
Communicate	Educate or talk down
Be likeable (remember that they may be stuck on a plane with you at some point)	Be arrogant
State your value proposition up front	Be vague about your technology or underestimate the importance of science
Come prepared with sufficient data, including backup slides	Deluge investors with facts
Enjoy yourself and let it show	Act desperate for funding
Keep the presentation within the allotted time slot	Act like you don't need money
Be realistic about valuations in the current market	List 'the company is undervalued' as a reason to invest
Make due diligence easy	Overprice your rounds so you can keep stepping up valuation
Realize that VCs are always thinking about exit strategy	Give investors a reason to turn you down

Slide 1: title

The purpose of the title or introductory slide is to briefly introduce your company without distracting from your spoken introduction. Cluttering this slide with unnecessary text, such as founding date or number of employees, forces the audience to read rather than pay attention to the speaker. To keep the audience's attention where it belongs, limit the content of this slide to your company logo and a tag line that quickly explains your business and, if possible, begins to differentiate your company from the competition.

A powerful spoken introduction is key to grabbing your audience's attention and setting the tempo for the rest of your presentation. To rouse your audience from the stupor that often sets in after sitting through numerous presentations, begin with a startling fact, famous quote or personal anecdote while your title slide is on the screen. If you remember one thing from this article, remember to be enthusiastic—your slide show is important, but ultimately, people buy from people, not from PowerPoint.

If you are not able to convey your enthusiasm about your company, it is difficult for investors to get excited about investing. When a VC sits down with a CEO who is obviously excited about his or her company, that enthusiasm (without being over the top) is infectious and makes a big difference in how the VC views the rest of the presentation (**Table 1**).

Slide 2: investment rationale and value proposition

The investment rationale is the single most important slide in your presentation—important enough to repeat three times.

According to the rule of communications math, $9 \times 1 = 0$, but $3 \times 3 = 1$. In other words, if you say nine messages one time each, your audience is unlikely to remember any of them, but if you say three messages three times each, your audience may remember one. This is why you begin and end with your investment rationale, which reiterates what you say during the body of the presentation. I like to call this the bookend approach.

An effective investment rationale slide explains why investors should give you money

Carin Canale is president at Porter Novelli Life Sciences, New York, New York, USA.
e-mail: ccanale@pnlifesciences.com

Table 2 The dos and don'ts of PowerPoint presentations

Do	Don't
Use only one topic per slide	Use sounds with your slide transitions
Limit the amount of text on each slide	Overdo the ALL CAPS, **bolded**, *italicized* or underlined text
Choose fonts and colors that are easy to read	Use too many different fonts in a presentation
Remember that PowerPoint is only an accessory to your presentation	Overuse special effects that steal center stage from the content of your presentation
Remember to spell check	Have technical difficulties—testing the presentation before the meeting is key

by listing five to seven bullet points that outline both your strengths and the direction of your presentation. Suggested topics for bullet points include core technology, product candidates, market opportunities, key partnerships or management strengths (**Table 2**).

Slide 3: business strategy

The business strategy slide answers the question on every potential investor's mind—how will your company make money?

VCs want to see up front how and when a company is going to make money, even if the company is early stage and profits are ten years away. There are thousands of great ideas, but only a limited number of them are supported by a sound business model.

Other subjects to address on your business strategy slide may include potential licensing opportunities to generate near-term revenue and manufacturing and commercialization strategies. You can also mention if and when you plan to partner with a larger company or how you plan to penetrate your target markets.

Slides 4–10: technology and/or products

Your technology and product slides make up the 'meat and potatoes' of your investor presentation—they are the central focus, but having too many will make your audience's eyes begin to glaze over. On the other hand, many investors are MDs or PhDs and like to see detailed data. The solution to managing the needs of your various audiences lies in backup slides.

Once you've given your presentation to a handful of investors, you tend to know which slides will lead to additional questions, so bring backup data slides to address these issues. One VC also recently advised me to "be up front about data that are less flattering, too, because I will find out about it, and if you don't tell me up front, how can I believe anything else you say?"

The order of your technology and products slides may vary depending on the stage of your company. Early-stage companies should begin by discussing the platform technology on which their products will be built, whereas later-stage companies should begin with the pipeline of products that provide a compelling investment opportunity. If your company has multiple products, beginning with a pipeline provides your audience with a quick overview and often elucidates how your products complement each other.

Technology slides should avoid jargon and provide a clear explanation of a company's science. Most investors agree that if you cannot explain your technology simply, you don't understand it. On the other hand, investors generally will not expect answers to very technical questions from CEOs or chief financial officers who do not have scientific backgrounds. It is perfectly acceptable to say you need to speak with your chief scientific officer before answering a question—just remember to follow up and provide the answer when you get it.

Product slides should address development status, disease indication, market size and recent data. A pet peeve of many investors, however, is a presentation that overinflates the market.

Choose data slides that are easy to comprehend even without your commentary to explain them, especially if you provide handouts of your presentation. In addition, forgo the typical scientific chart and graph titles in favor of more active titles that help to interpret the results they illustrate, such as 'Product X decreases tumor progression' or 'Product X delays need for chemotherapy'.

Slide 11: intellectual property

Contrary to popular belief, your intellectual property slide should not list every patent you've ever received. Instead, outline a patent strategy that shows you can't be touched and list only key patents.

Slide 12: management team and advisors

Another frequently misused slide is the one for the management team and advisors. This slide not only should list the names of top executives and advisors, but also should include a few choice words explaining what each brings to the table, such as 'Former VP of oncology drug development at GlaxoSmithKline' or 'Closed three big pharma deals and secured $50M in equity financing'.

Another investor pet peeve is listing advisors who are not actively working with the company. Don't talk about an advisor who is not currently involved with the company, because VCs don't want to call him or her only to hear you just paid for the name. On the other hand, if you have a phenomenal chief medical officer or world-renowned chairman, don't hesitate to bring that person to meetings with you.

Slide 13–14: recent achievements and upcoming milestones

The recent achievements slide should illustrate—especially to those audiences who have seen your presentation in the past—that you deliver on your promises. Consider including data recently presented or published, financings, key executives added, clinical trial milestones (initiation, enrollment, completion, etc.) and other significant achievements.

If there are any perceived weaknesses about your company, your anticipated milestones slide should explain how you plan to address them. For example, if your management team lacks breadth, list milestones such as 'Hire key finance and business development executives'. Just be careful to list only milestones you expect to deliver on—and be sure to include a few near-term items that are already 'in the bag'. This can go a long way toward establishing credibility.

Slide 15: investment rationale

A good presentation begins with a summary of what you are about to say and ends with a summary of what you've just said. Repeating your investment rationale slide at the end of your presentation also ensures that your audience is left with the key message points you are trying to convey. If you're lucky, according to the rule of communications math, they might even remember one point.

Conclusion

Following the basic format outlined above will not guarantee financing, but it will help set you apart from the crowd and leave a lasting impression with the VCs. Most importantly, remember to finish your slide presentation well in advance and take time to practice. As stressed above, enthusiasm is contagious—your slide show is important, but ultimately, PowerPoint is only a tool; you are the key. Your presentation is your story and all audiences appreciate a good story. Have fun and good luck.

Pharma's role is not to bankroll biotech

To the Editor:

The editorial in the February issue entitled 'The worst of times, the best of times'[1] is well meaning and timely but misunderstands the nature of big pharma's relationship with small biotechs.

Your hypothesis—"big pharma should be more proactively investing in cash-hungry biotech companies"—is supported by data showing large companies have cash reserves plus two impossible-to-prove assertions. First, that we underestimate the "promising products" from "undervalued" biotechs; second, that biotechs are our "drug discovery engine."

Let's start with the word "should" in your hypothesis. Those of us who manage R&D investments prefer the word "must." We must invest our shareholders' funds in areas of unmet medical need. We must consider the feasibility and/or practicality of the science and likelihood of success. We must have evidence that payers will value our experimental medicines.

All this acknowledges a simple truth of our industry—there is no shortage of good ideas. Instead, we are exhilarated by the enormous number of opportunities—from within our own laboratories and from outside. Success is picking and nurturing those few with real potential. At Pfizer (New York), our choices are guided by the criteria above plus a five-point strategy that includes the directive "pursue the best external science."

As president of global research and development at Pfizer, I oversee an extensive pipeline. The majority of projects in that pipeline have come from our own laboratories, but I gladly acknowledge those discovered elsewhere. Our drug discovery engine is, in fact, a broad federation of in-house and external science. We are doing everything possible to maintain that diversity. Together with our Biotherapeutics and Bioinnovation Center, we fund academic work, incubate startups, collaborate on early science and partner in development. Two examples illustrate how, sometimes, we take on all the risk.

Sutent (sunitinib malate) is Pfizer's oral multi-kinase inhibitor indicated for the treatment of advanced renal cell carcinoma and gastrointestinal stromal tumor (GIST). Other indications are under investigation. It was discovered by the biotech company Sugen (formerly of San Francisco, before acquisition by Pfizer in 2003) but was not that company's first choice for development. The medicine's success is a tribute to Sugen's chemistry, plus significant scientific, medical and other investments from Pharmacia (Kalamazoo, MI, USA), then Pfizer.

Acquired as part of Pfizer's 2006 purchase of Rinat (S. San Francisco, CA, USA), tanezumab is a humanized monoclonal antibody designed to have high specificity and affinity for nerve growth factor. Clinical efficacy was recently demonstrated in the treatment of osteoarthritis in phase 2 trials, and phase 3 clinical studies were initiated in November last year. Tanezumab is poised to be the first biologic agent approved specifically for the treatment of pain, and it may transform the way severe, unremitting chronic low back pain is treated. Pfizer essentially assumed all of the development risk with this compound.

These two anecdotes, plus the thousands of smaller partnering deals, point to our keen appreciation for benefit sharing and financial risks. Our knowledge of biotech is considerable, we listen carefully to our external advisors and our sensitivity is based on decades of partnering with smaller biotechs and technology companies.

On behalf of our shareholders, we are enthusiastic small biotech investors but we cannot, and should not, adopt all the risks now owned by the broader financial community.

COMPETING INTERESTS STATEMENT

The author declares competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturebiotechnology/.

Martin Mackay

Pfizer Inc., Global Research & Development, New London, Connecticut, USA.
e-mail: Martin.Mackay@Pfizer.com

1. Anonymous. *Nat. Biotechnol.* **27**, 101 (2009).

Conflating MTAs and patents

To the Editor:

It is unfortunate that the paper by Zhen Lei, Rakhi Juneja and Brian D. Wright entitled "Patents versus patenting: implications of intellectual property protection for biological research" in your January issue[1] obscures an important result with the red herring of "patents are bad for research." Indeed, the piece records that a cohort of agricultural scientists from leading research schools have a subjective belief that patenting has a negative affect on research. Paradoxically, however, respondents reported that they routinely ignore the existence of patent protection for research tools. More than 90% of respondents report that they "have never checked whether a tool that they might need in planned research is patented." The reason, according to the scientists, is that most think they won't be sued.

Upon reading the article, it is clear that the scientists polled are woefully misinformed about the difference between patents and intellectual property (IP), and that most of their responses are self-serving and reflect the cultural differences between academics and industry, with university technology transfer professionals being caught in the middle. The issue is not patents, but rather material transfer agreements (MTAs), private contracts between research universities that govern the disposition of tangible research materials. There are many and significant differences between

patent protection and MTAs. For example, patents promote disclosure, whereas MTAs typically require continued confidentiality. MTAs are exactly that: agreements concerning the transfer of materials. This means they are limited to tangible items that can be transferred and exclude IP, such as know-how, trade secrets and methods (indeed, their tangibility makes them more akin to personal property than IP). Patent rights are exhausted by a sale, whereas with MTAs the granting institution typically retains ownership of the transferred materials and requires either their return or certification that they have been destroyed after the term of the agreement has expired. In addition, although patents are governed by federal statute, and are encumbered with protections against improper use, MTAs are private contracts between the parties, governed by state common law that typically permits any behavior not in direct contravention of criminal or other statutes (that is, contract law is much more permissive than patent law).

Thus, the actual conclusions of the paper are not related to the effects of patenting on academic research at all. Rather, the authors report that institutionally mandated MTAs delay research, and these MTAs put "sand in the wheels" of an otherwise "lively system of interdisciplinary exchanges" of research materials. I do not doubt the researcher respondents feel this way; however, the disparity between these results and the results of several other academic reports (which argue that IP protection has a negligible effect on academic research) should raise a few questions about the nature of the study and the elicited responses. Academic researchers are focused, ambitious (and some would say even egotistical) people used to having their own way; these traits are perhaps necessary for them to have the temerity to believe they can make sense of a complex world, and are certainly an expected consequence for individuals having the intelligence of most academic researchers. The law presents them with another, different set of rules and a logical structure that differs from science. Particularly in view of the power differential between tenured professors and the staff of most university technology transfer offices, the scientists frequently believe they can ignore the rules (see their disdain for potential patent infringement reported in the paper), or if 'forced' to comply believe that it must have a negative effect on the only thing they are interested in, getting their research done as timely as possible (because there are usually other researchers actively engaged in their area).

Indeed, rather than patenting or other IP protections, academic competition may be the greatest impediment to the 'free exchange' of research materials and information. As the study authors admit, "[l]ong before the proliferation of IP protection, scientists were often secretive and uncooperative in their interaction with competitors (Hagstrom, W.O., *Am. Sociol. Rev.* **39**, 1–18, 1974)," and "[Respondents] anticipate moderate degrees of difficulty ["3.2 on a 5-point scale"] in getting tools from rivals…."

But recognizing these nuances of the problem is not as 'sexy' as pitching the results as being "contrary" to the "developing consensus" that patents have not had a negative effect on university research. Although the authors believe that there is an advantage to obtaining "direct" results of the effects of "IP protection" from the researchers, an uncritical acceptance of the responses and a failure to appreciate the important distinctions between MTAs and patents (which promote disclosure and hence academic cooperation and the free flow of information) leads them to conclude that IP protection impedes academic freedom and stifles research. From the responses reported in this paper, nothing could be further from the truth, and failing to address or even simply report that does little to illuminate an important issue for US patent policy.

Kevin E Noonan

McDonnell, Boehnen, Hulbert & Berghoff, LLP, Chicago, Illinois, USA.
e-mail: noonan@mbhb.com

1. Lei, Z., Juneja, R. & Wright, B.D. *Nat. Biotechnol.* **27**, 36–40 (2009).

Zhen Lei and Brian D. Wright reply:

To a reader unfamiliar with intellectual property (IP), Noonan's thesis might well be persuasive. Researchers have problems with material transfer agreements (MTAs), not patents. MTAs are different from patents, and more "akin to personal property than IP." Indeed, they are "limited to tangible items that can be transferred and exclude IP, such as know-how, trade secrets and methods." Noonan implies that MTAs are not used in the transfer of IP, so scientists surveyed in our paper[1] are "woefully misinformed" when they attribute problems with MTAs to the recent proliferation of patents and other IP.

Scientists who rely on the counsel of attorneys or Office of Technology Transfer personnel, or draw on their own experience of patenting tangible research tools, understand that patentable compositions of matter, including those that are research tools, are IP. They also understand that their Offices of Technology Transfer have, since the 1980 Bayh-Dole Act, taken a greater interest in patenting and other means of IP protection, and urged scientists to use MTAs in sending research tools to others or receiving materials from peers.

For example, the relevant University of California, Los Angeles website[2] advises: "The purpose of the MTA is to protect the intellectual and other property rights of the provider while permitting research with the material to proceed." Furthermore, "If the material is not yet patented (or, publicly disclosed) and of possible commercial value, a material transfer agreement with secrecy provisions may be required." For scientists on the research frontier, the tools they want to exchange, often unpatented at the time of transfer, may be protected by MTAs as part of a strategy for preserving rights to royalties, and other benefits from patents or other IP related to inventions arising from the materials transferred. Another aspect is that MTAs might restrict use of materials in ways that go beyond what a patent would protect.

Since 1980, patenting by academic institutions has greatly increased. MTAs on materials sent from academia and industry "are often associated with having patent rights to the material in question"[3]. Scientists surveyed in the United States and Japan by the American Association for the Advancement of Science (AAAS; Washington, DC, USA) report that ~30% of the patented technology they acquired was transferred via MTAs; a substantially smaller portion was acquired by licensing[4]. It is not surprising, then, that the scientists we surveyed perceive a connection between the surge in patenting and the proliferation of MTAs on transferred tools.

Indeed, the connection between patenting and MTAs is evident in the behavior of our own respondents. When the nonpatentees among them provided tools to academic peers, they used MTAs in only 12% of the cases, whereas formal contracts (predominantly MTAs) covered 34% of such transfers by patentees. (Noonan will surely concede that these patentees should be familiar with the distinctions among patents, MTAs and other types of IP. Nevertheless, patentees agree with their peers on the net effects of intellectual protection on research.)

Noonan conjectures that the greatest impediment to tool exchange might be academic

competition. However, our respondents anticipate only moderate difficulty with rivalry, and Noonan's conjecture misses the nuance that our scientists report no recent change in such competition, whereas they have seen their problems with tool exchanges increase. Moreover, in none of the 17 cases covered in our follow-up interviews was academic competition the dominant factor impeding access to a research tool.

Although we do agree with Noonan that there is a need to distinguish between 'patents' and the broader term 'intellectual property', unfortunately, he honors this distinction in the breech. Contrary to Noonan's claim, our results do not conflict with other academic reports. These focus on the direct effects of existing patents. For example, only questions 48 A–F of the four AAAS reports[5–8] ask specifically about IP protection as such. The responses, for large multidisciplinary samples of scientists in four countries, are in general remarkably supportive of our findings, though they are not discussed in any of these reports.

Thus, the paradox encountered by Noonan is resolved. Academic scientists are not greatly restricted by the need to avoid infringing existing patents because they are rarely aware of such patents and the tools they use are often too new to be patented. Even so, their work is, overall, affected indirectly by the institutional promotion of the use of MTAs, induced largely by the proliferation of patenting in academia and in industry, and this effect outweighs any incentive-related effects of patenting.

Finally, we have tried to avoid hyperbole and oversimplification in discussing this complex issue. We believe that patenting of research tools rarely 'stifles' a research project. Rather, proliferation of patenting and other IP protection of research tools has led to an increase in the use of MTAs. Resulting difficulties with research tool exchanges make the research progress of the agricultural biologists we surveyed sufficiently slower or more difficult that they believe that the costs of IP protection outweigh the benefits.

We find no reason to believe that these scientists are misinformed about these issues. It is possible that scientists fail to perceive some important social benefits from patenting their research tools. If scientists' views are surprising to some who have confused the effects of existing patents with the full implications for research of the proliferation of IP, then they are all the more valuable as a contribution to an ongoing debate.

1. Lei, Z., Juneja, R. & Wright, B.D. *Nat. Biotechnol.* **27**, 36–40 (2009).
2. <http://www.research.ucla.edu/mta/#mta>
3. Rodriguez, V. *Nat. Biotechnol.* **23**, 489–491 (2005).
4. Hansen, S.A. *International Intellectual Property Experiences: A Report of Four Countries* (The Project on Science and Intellectual Property in the Public Interest (SIPPI), American Association for the Advancement of Science (AAAS), Washington, DC, 2007). <http://sippi.aaas.org/Pubs/SIPPI_Four_Country_Report.pdf>
5. Hansen, S.A., Kisielewski, M.R. & Asher, J.L. *Intellectual Property Experiences in the United States Scientific Community* (The Project on SIPPI, AAAS, Washington, DC, 2007). <http://sippi.aaas.org/Pubs/SIPPI_US_IP_Survey.pdf>
6. Kisielewski, M.R. Asher, J.L. & Hansen, S.A. *Intellectual Property Experiences in the United Kingdom Scientific Community* (The Project on SIPPI, AAAS, Washington, DC, 2007). <http://sippi.aaas.org/Pubs/SIPPI_UK_IP_Survey.pdf>
7. Westerburg, S., Asher, J.L., Kisielewski, M.R. & Hansen, S.A. *Intellectual Property Experiences in the German Scientific Community* (The Project on SIPPI, AAAS, Washington, DC, 2007). <http://sippi.aaas.org/Pubs/SIPPI_Germany_IP_Survey.pdf>
8. Walsh, J.P. & Huang, H.I. *Research Tool Access in the Age of the IP Society: Results from a Survey of Japanese Scientists* (The Project on SIPPI, AAAS, Washington, DC, 2007). <http://sippi.aaas.org/Pubs/SIPPI_Japan_IP_Survey.pdf>

First GM trial in Belgium since 2002

To the Editor:

A news article in your February issue[1] reported that GM poplars developed by the group of Wout Boerjan at the Flanders Institute of Biotechnology (VIB) in Ghent were to move to the Netherlands to go on trial there. I am happy to report that VIB finally succeeded in getting an authorization for the trial in Belgium and does not have to move abroad.

The application in Belgium was first refused in May 2008, even though the Belgian Biosafety Advisory Council and the regional Flanders minister of the environment had both given their positive advice. VIB took legal action at the Council of State (the highest Belgian court) and made a few rounds of negotiations to overturn the negative decision and finally get the authorization in mid-February 2009.

The authorization is a landmark in the genetically modified organism field trial history in Belgium. It is the first field trial in Belgium since 2002. From 1987 to 2002 Belgium had a flourishing field trial culture reflecting the country's advanced research in plant biotech. In 1983, researchers in Ghent led by Marc Van Montagu and Jef Schell were the first to develop a genetically engineered plant. The trial in 1987 was one of the first in the world, but after 2002, the number of field trials dropped down to zero as the result of regulatory uncertainty surrounding the implementation of the 2001/18 EU directive on the deliberate release of gentically modified organisms. Laboratory research on plant biotech, however, has always kept up its pace.

Even though VIB has successfully pursued a field trial permit in The Netherlands as well, it will not start a trial there in the near future. It commenced planting of its trees last month on a field trial plot in Ghent. The plot is close to the research facilities and also close to the biofuels pilot plant, which is being set up in the port of Ghent. In trees themselves lignin biosynthesis is suppressed leading to trees with about 20% less liginin and 17% more cellulose per gram of wood. This makes them more suitable for bioethanol production. Wood from these trees grown in the greenhouse produces up to 50% more bioethanol than ordinary poplar trees. The field trial is the ultimate test to see whether wood produced under real-life conditions—seasons, stormy weather and a marginal soil—is also able to produce ethanol in a much more efficient way. VIB expects to have its first results from the trial in 2012.

René Custers

VIB, Ghent, Belgium.
e-mail: rene.custers@vib.be

1. Birch, H. *Nat. Biotechnol.* **27**, 107 (2009).

A European perspective on immunogenicity evaluation

To the Editor:

We read with interest the Perspective appearing in the May 2007 issue entitled "A risk-based bioanalytical strategy for the assessment of antibody immune responses against biological drugs"[1]. This article correctly states that the recently published European guideline on "Immunogenicity Assessment of Biotechnology-derived Therapeutic Proteins"[2] does not give specific recommendations on a method for risk assessment or on the extent of the requisite characterization of anti-drug antibodies (ADAs). As European regulators having been involved in the drafting and finalization of this European guideline on immunogenicity issued by the Committee for Medicinal Products for Human Use (CHMP) of the European Medicines Agency (EMEA; London), we would like to comment on some of the ideas put forward in the publication by Shankar *et al.* and to contribute to this important topic some further reflections that need to be considered from a regulatory perspective in the context of the overall benefit-risk assessment that is the regulators' central scientific task.

In September 2007, the EMEA organized a workshop on immunogenicity assessment, giving various stakeholders the opportunity to comment and critically discuss the European draft guideline[3]. From the comments received before and during that meeting, it became clear that an ideal guideline for industry would be a detailed 'fit-to-purpose cooking recipe' on how to plan, design and conduct immunogenicity assessment for novel biotech-derived medicinal products. It became clear also that many experts in industry would like to have a standard algorithm defining the risk-based approach.

It was rather obvious that the interpretation of the guideline text was different from the intention of the regulators drafting the guideline. The intention of this particular guideline is to give general key principles and strategy for assessing immunogenicity. The EMEA/CHMP guideline advocates a multidisciplinary risk- and science-based approach to immunogenicity, as experience both with the evaluation of various 'marketing authorisation' dossiers and 'scientific advice' procedures suggests that many applications lack this kind of comprehensive and prospective approach. In fact, European regulators wish to maintain flexibility of their guidance and welcome novel approaches in this rapidly progressing field.

Biologicals have to be seen as individuals. Thus, there is hardly an ideal fit-to-purpose recipe for immunogenicity evaluation.

A risk-based approach helps the developer to justify the selected immunogenicity program. The concept for a risk-based approach, as proposed by Shankar and colleagues[1], is to distinguish among biological drugs with lower risk (e.g., antibody drugs designed to bind and inhibit proteins like tumor necrosis factor-α), medium risk (e.g., β-interferons or agonistic antibodies) and high risk (e.g., erythropoietin). On the basis of their classification, Shankar *et al.*[1] propose to adapt the regulatory requirements accordingly as regards both the frequency of sampling during clinical trials and the extent of characterization of ADAs (e.g., the need to evaluate their neutralizing potential).

Although the perceived risk is likely to affect the choice of evaluation approach, in our opinion, the use of a standardized algorithm as a general tool to define the strategy for immunogenicity evaluation of biotechnological medicinal products is not appropriate now because it might, in some cases, lead to regulatory concerns. There are various reasons for this; for instance, the more explicit the recommendations given (e.g., classify all anti-cytokine antibodies as lower-risk drugs), the higher the chance to miss cases that have unexpected features that arise, for example, from manufacturing and quality-related aspects. To illustrate this, a biological drug classified as lower risk on the basis of the structure of the active substance or mechanism of action may require a different strategy for immunogenicity evaluation if produced by an entirely novel expression system, such as transgenic plants or transgenic animals. Similarly, an active substance sourced from a new type of expression system is likely to pose a higher risk than the same active substance sourced, manufactured and controlled with more established approaches and methods.

Another important aspect to consider is the impact on overall benefit-risk assessment. For

example, infusion reactions for 'lower risk' drugs can be clinically manageable and are usually of less concern than neutralizing antibodies cross-reacting with endogenous proteins. Nevertheless, ADAs for low-risk drugs can expose patients to clinical conditions that affect the overall benefit-risk assessment that is the central regulatory task before approval. The same holds true for a loss of efficacy by neutralizing antibodies. The situation is further complicated by the finding of different immune responses to a given product in different diseases and patient populations[4]. In other words, although overall a biological drug might be a lower-risk compound, the individual clinical sequelae are of high importance for the individual patient and have an impact on overall regulatory benefit-risk analysis.

Some explicit recommendations by Shankar *et al.*[1], such as the characterization of ADA positives to be explored only when considered necessary or decisions per clinical trial within a program, also pose the danger that data generated in a strategy for a predefined 'medically low risk' drug turn out to be insufficient for a proper benefit-risk assessment and thus do not meet regulatory requirements at the time of approval.

Thus, although many of the ideas developed by Shankar *et al.* have the merit of focusing on a risk-based approach and are valid and indeed interesting, we feel that their scheme of risk-based classification of products serve as a useful starting point for reflection at the time of planning immunogenicity studies as part of clinical trials; however, they may not in all cases be sufficient to support the benefit-risk evaluation required at the time of licensing of the product. Companies still need to justify their approach when filing a marketing authorization application in the European Union. We believe the European approach to immunogenicity, as presented in the final guideline document, retains a good balance by providing guidance on the conceptual planning of an immunogenicity evaluation on one hand, but being sufficiently open-minded to allow the flexibility needed for 'individual' biological drugs, on the other hand.

DISCLAIMER

C.K.S. is chairman of the CHMP Working Party on Similar Biological (Biosimilar) Medicinal Products Working Party (BMWP) and P.K. is a former chairman of BMWP. The views expressed in this article are the personal views of the authors and may not be understood or quoted as being made on behalf of or reflecting the position of the EMEA or one of its committees or working parties.

Christian K Schneider[1,3], Marisa Papaluca[2] & Pekka Kurki[4]

[1]Paul-Ehrlich-Institut, Federal Agency for Sera and Vaccines, Langen, Germany. [2]European Medicines Agency, London, UK. [3]Twincore Centre for Experimental and Clinical Infection Research, Hannover, Germany. [4]National Agency for Medicines, Helsinki, Finland.
e-mail: schci@pei.de

1. Shankar, G., Pendley, C. & Stein, K.E. *Nat. Biotechnol.* **25**, 555–561 (2007).
2. European Medicines Agency. Guideline on Immunogenicity Assessment of Biotechnology-derived Therapeutic Proteins EMEA/CHMP/BMWP/14327/2006. <http://www.emea.europa.eu/htms/human/humanguidelines/multidiscipline.htm>
3. BMWP/BWP Workshop on Immunogenicity Assessment of Therapeutic Proteins, EMEA, London, September 4, 2007. <http://www.emea.europa.eu/meetings/conferences/4sep07.htm>
4. European Medicines Agency. European Public Assessment Report (EPAR) for Remicade. <http://www.emea.europa.eu/humandocs/Humans/EPAR/remicade/remicade.htm>

Reflect: augmented browsing for the life scientist

To the Editor:

Anyone who regularly reads life science literature often comes across names of genes, proteins or small molecules that they would like to know more about. To make this process easier, we have developed a new, free service called Reflect (http://reflect.ws) that can be installed as a plug-in to web browsers, such as Firefox or Internet Explorer. Reflect tags gene, protein and small-molecule names in any web page, typically within a few seconds and without affecting document layout. Clicking on a tagged gene or protein name opens a popup showing a concise summary that includes synonyms, database identifiers, sequence, domains, three-dimensional structure, interaction partners, subcellular location and related literature. Clicking on a tagged small-molecule name opens a popup showing two-dimensional structure and interaction partners. The popups also allow navigation to commonly used databases. In the future, we plan to add further entity types to Reflect, including those outside the life sciences.

As science uncovers the intricate interconnections within biological systems, many life scientists constantly come across unfamiliar biochemical entities (e.g., genes, proteins or small molecules) that were previously not known to be relevant to a given field, but where today's literature shows an important, new connection. For such cases, it is clearly valuable to systematically tag all scientific entities in a publication, thus helping the reader to navigate to more specific information about any entity of interest. Such tags can help the reader to comprehend scientific content more rapidly and completely. Even when an entity is already familiar to a reader, it can be valuable to have quick access to commonly used source data entries; for example, protein sequences or two-dimensional structures of small molecules.

In spite of the clear value of systematically tagging scientific entities, only a small fraction of the main scientific publishers currently offer such tags on their web content. Some publishers are beginning to explore the option of adding tags as part of the publication process[1]; however, enforcing, validating and updating these tags creates additional work for publishers and authors.

The task of accurately tagging biochemical entities automatically is very challenging; this task has been the subject of intense research efforts that has led to significant improvements in accuracy[2]. These automated methods have been used to develop a wide variety of text mining applications and services, many of which are designed to provide sophisticated search, analysis and presentation capabilities[3]. However, a few text mining services have been designed to appeal to the broader life science community; for example, iHOP[4] provides simple search, navigation and presentation of Medline abstracts with systematically tagged gene and protein names.

Tagging a scientific entity is only half the story: the other half is the information that is accessed when the user clicks on a tag. In the past, entity tags were almost always simple hyperlinks to web pages showing source data entries. Increasingly, however, entity tags are not hyperlinks but scripts that create a small popup window (typically with Javascript). A key advantage of using popups is that users can see basic information about an entity without having to navigate away from the current web page. If needed, hyperlinks to more detailed information can be provided on the popup.

An emerging trend is to augment normal web browsing by using plug-ins, such as Greasemonkey (http://greasespot.net/), that let end-users modify the appearance of web pages while browsing. We believe that such augmented browsing tools will soon have an important impact on how scientists read literature on the web. For example, one such tool, ChemGM[5], lets end-users tag small-molecule names in any web page; clicking on a tagged small molecule opens a popup that shows the two-dimensional

structure. Tagging is done by sending the page to a remote server, and the total time taken is typically about one minute for a five-page document. Another tool, Concept Web Linker (http://conceptweblinker.wikiprofessional.org/), has a broader scope: it tags a range of entities, such as genes, chemicals and diseases, again typically within about one minute. However, the Concept Web Linker popups show less specific information, giving only a short text description for each entity; to reach more specific information, such as protein sequences, the user needs to navigate through a series of web pages, in some cases browsing complex ontologies. A related system, Cohse[6], has even broader scope—it enables users to choose many different ontologies, including those outside the life sciences. Currently, however, the publicly accessible versions of Cohse provide only very limited functionality and using the life-science ontologies provided does not allow direct navigation to specific information, such as sequences.

We designed Reflect to be an augmented browsing tool that would be broadly useful to life scientists, and would address the limitations of the above tools. A primary goal of Reflect was to enable the user to navigate directly from a gene or protein name to a specific sequence. A second goal was to be able to tag a typical web page in a few seconds. A third goal was to provide entity popups that give a concise summary of the most important features of the entities, as well as direct hyperlinks to commonly used source data entries (**Fig. 1** and **Supplementary Methods** online). Finally, Reflect was designed with a strong focus on ease of installation and on usability.

Reflect can be used directly from http://reflect.ws/ by typing or pasting in a URL. In this case, the Reflect server retrieves the HTML document, tags it and returns the tagged version to the user's browser. Note that this will work only for URLs that are publicly accessible.

A more convenient way to use Reflect is to install it as a plug-in to Firefox or Internet Explorer. In this case, the HTML document is retrieved by the user's browser, then sent to the Reflect server, tagged and returned to the browser. Thus, with the plug-in, users can 'Reflect' any page that they can access.

The Reflect server at the European Molecular Biology Laboratory keeps in RAM (random-access memory) a large dictionary with names and synonyms for 4.3 million small molecules, and for 1.5 million proteins from 373 organisms. When tagging an HTML document, the server finds all occurrences of these synonyms and returns a slightly modified version of the HTML document to the user's browser—the only difference is that all matching protein, gene and small-molecule names are now tagged and

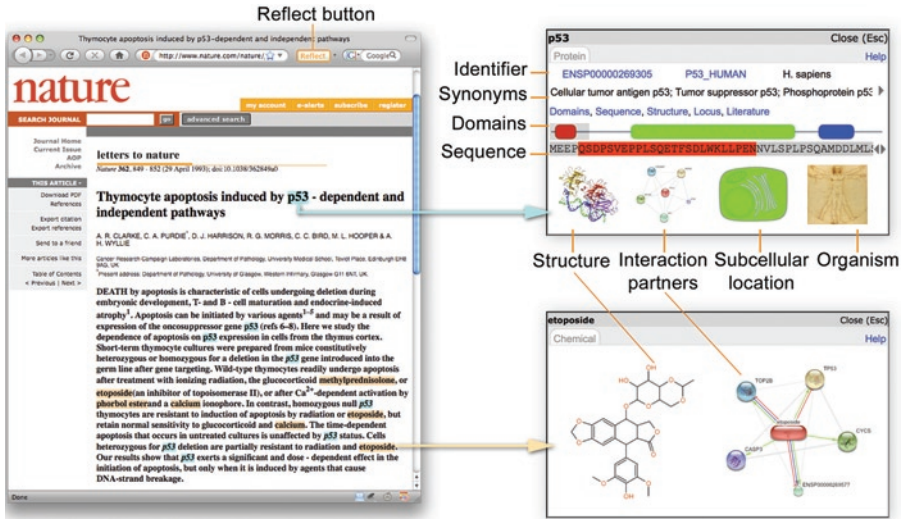


Figure 1 The Reflect button can be installed in the Firefox or Internet Explorer web browsers. Clicking the Reflect button tags protein and gene names (blue highlighting), and small molecules (orange highlighting) in any web page. Clicking on a highlighted name opens a small popup showing a concise summary of important features of the entity, and provides access to related information (**Supplementary Methods**).

highlighted. Tagging a document usually takes much less time than uploading and downloading it; thus, the time taken for the entire process (upload, tag and download) depends almost exclusively on the speed of the user's internet connection. With standard broadband, the entire process usually takes from one to five seconds for a five-page document (**Supplementary Methods**).

Clicking on a tagged small-molecule name opens a summary popup (**Fig. 1**, bottom right) that shows two-dimensional structures from PubChem[7] and interaction partners from STITCH[8]. Clicking on a tagged protein or gene name opens a popup (**Fig. 1**, top right) that shows synonyms, the complete amino acid sequence of the longest transcript, domains from the SMART[9] database, a representative three-dimensional structure from PDBsum[10], principal interaction partners from STITCH[8], known subcellular location and an image of the organism. Most of these features on the popup are hyperlinked to related database entries. The popup also has hyperlinks to the corresponding gene entry and to related Medline abstracts in iHOP[4]. Dragging the mouse on the domain graphical view scrolls through the sequence, and hovering over a domain causes the domain name to appear in a tool tip.

When a tagged name is ambiguous, the popup shows all possible matches and allows the user to disambiguate the name by choosing which of the possibilities is most appropriate. Currently, three levels of ambiguity are shown. First, a name may match both a protein and a small molecule; in this case, Reflect shows both possibilities on separate tabs. Second, a name may match to several genes within the same organism; here, Reflect shows all matching genes in a pull-down menu. And third, for gene and protein names, it is often ambiguous which organism is intended in the HTML document; to address this, Reflect shows a list of possible organisms derived from the default organism (which is initially set to human, but can be changed using the Firefox plug-in) plus organisms mentioned in the document. In the near future, we also plan to show a fourth level of ambiguity, where users will be able to select splice variants for each gene.

Any automated method for recognizing biochemical entity names will make some errors: some false positive matches will arise due to overlap with commonly used words or acronyms, and false negatives will arise due to incompleteness of the tagging dictionary. To assess the accuracy of Reflect, we tested it against the BioCreative[11] benchmarks. Compared with 15 other tools for automated entity recognition that were assessed in BioCreative, Reflect ranked second best (91% *F*-score) using the *Saccharomyces cerevisiae* benchmark and had median performance (66% *F*-score) using the *Drosophila melanogaster* benchmark. We consider these to be quite good results because, unlike the other tools tested against these benchmarks, Reflect was designed to optimize speed rather than accuracy.

In the near future, we plan to enable community-based, collaborative editing for some of the information in Reflect popup, especially the synonym lists. These and other planned extensions will enable the user community to improve Reflect by correcting false-negative and false-positive matches. We plan to add further

entity types (e.g., diseases, pathways and organisms), and eventually to add entity types beyond the life sciences; we designed Reflect to be an extendible platform, and we welcome collaboration proposals for adding further entity types. In addition, we welcome proposals from publishers and data providers interested in programmatic access to Reflect. With such access, end-users can use 'Reflected' content without needing to install a browser plug-in.

In summary, Reflect creates a view of the web tailored for the life scientist, that is, with systematic tagging of biochemical entities, and easy access to more detailed information. Reflect is already being used by thousands of researchers, and we have received much positive feedback regarding Reflect's usefulness and ease of use. In addition, just before publication of this correspondence, Reflect was awarded first prize in the Elsevier Grand Challenge, a contest for tools that improve the way scientific information is communicated. Thus, we believe that Reflect can be a valuable tool for researchers, teachers, students and anyone who reads life science literature on the web. We further predict that in the near future tools such as Reflect will change dramatically how scientists use the web.

Note: Supplementary information is available on the Nature Biotechnology website.

ACKNOWLEDGMENTS

Many thanks to Philippe Julien for the subcellular location viewer.

Evangelos Pafilis[1,3], Seán I O'Donoghue[1,3], Lars J Jensen[1–3], Heiko Horn[1], Michael Kuhn[1], Nigel P Brown[1] & Reinhard Schneider[1]

[1]European Molecular Biology Laboratory, Heidelberg, Germany. [2]NNF Center for Protein Research, University of Copenhagen, Denmark. [3]These authors contributed equally.
e-mail: contact@reflect.ws

1. Ceol, A., Chatr-Aryamontri, A., Licata, L. & Cesareni, G. *FEBS Lett.* **582**, 1171–1177 (2008).
2. Smith, L. *et al. Genome Biol.* **9** Suppl 2, S2 (2008).
3. Krallinger, M., Valencia, A. & Hirschman, L. *Genome Biol.* **9** Suppl 2, S8 (2008).
4. Hoffmann, R. & Valencia, A. *Nat. Genet.* **36**, 664 (2004).
5. Willighagen, E.L. *et al. BMC Bioinformatics* **8**, 487 (2007).
6. Bechhofer, S.K., Stevens, R.D. & Lord, P.W. *Pac. Symp. Biocomput.* **10**, 79–90 (2005).
7. Wheeler, D.L. *et al. Nucleic Acids Res.* **36**, D13–D21 (2008).
8. Kuhn, M., von Mering, C., Campillos, M., Jensen, L.J. & Bork, P. *Nucleic Acids Res.* **36**, D684–D688 (2008).
9. Letunic, I. *et al. Nucleic Acids Res.* **34**, D257–D260 (2006).
10. Laskowski, R.A. *Nucleic Acids Res.* **29**, 221–222 (2001).
11. Hirschman, L., Colosimo, M., Morgan, A. & Yeh, A. *BMC Bioinformatics* **6** Suppl 1, S11 (2005).

Extrapolating from sequence—the 2009 H1N1 'swine' influenza virus

To the Editor:

The recent incidence and spread in humans of the 'swine flu' influenza A virus has raised global concerns regarding its virulence and pandemic potential. The main cause of the so-called swine flu has been identified as human infection by influenza A viruses of a new H1N1 (hemagglutinin 1, neuraminidase 1) subtype, or '2009 H1N1 strain'. The first cases of human infection were reported in April in the Mexican town of La Gloria in Veracruz; soon after, reported infections occurred in areas of southern California and Texas. Several recent studies have focused on the necessary determinants for human adaptation and efficient human-to-human transmission of the H1N1 influenza A viruses[1–9]. Here, using a representative 2009 H1N1 strain as our starting point, we offer a perspective on the likely human adaptation and transmissibility of 2009 H1N1 viruses.

At the time when this sequence analysis was performed, partial or complete sequences were available from 38 different human isolates of the 2009 H1N1 virus. These sequences were obtained from GISAID (Global Initiative on Sharing Avian Influenza Data; http://platform.gisaid.org/) and the NCBI Influenza Virus Resource (National Center for Biotechnology Information; http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html). Comparison of the amino acid sequences between the 38 isolates showed some intragenic differences: seven amino acid positions in HA (hemagglutinin), one in M1 (matrix 1), two in M2 (matrix 2), four in NA (neuraminidase), three in NP (nucleoprotein), two in PA and two in PB2 (both of which encode subunits of viral RNA polymerase). Given the few intragenic variations among the 38 isolates available at the time of this study, we use /California/04/2009 (Cal0409) as a representative 2009 H1N1 virus strain for further analysis. The top ranking hits of the BLAST search using the individual Cal0409 genes are shown in **Supplementary Table 1** online.

Comparison of Cal0409 HA with the HA consensus sequences for human-adapted H1N1, avian-adapted H1N1 and swine-adapted H1N1 reveals important substitutions in positions 100–300, where the glycan receptor-binding sites and antigenic loops are located (**Supplementary Table 2** online). Notably, the Cal0409 HA possesses the signature amino acids Asp190 and Asp225 that have been shown to play a key role in conferring specificity to the human α2-6 sialylated glycan receptors[1]. We also observe amino acid substitutions that are unique to Cal0409 HA and have not been observed in previous human H1N1 HAs. These include substitutions at sequence positions 74, 131, 145, 208, 219, 261, 263, 264, 305, 317, 368, 377 and 530. Among these residue positions, 131 and 145 are proximal to the glycan-binding site.

To determine the possible effect of these mutations on the glycan-binding properties of HA, we constructed homology-based structural complexes of Cal0409 with representative α2-3 and α2-6 sialylated oligosaccharides, as described earlier[8] (**Fig. 1**). The construction of theoretical HA-glycan structural complexes previously[8] allowed us to provide a structural rationale for how specific amino acid mutations within the 1918 H1N1 HA can dramatically alter its relative α2-3/α2-6 binding affinity. Referencing these previous efforts, we determined the potential glycan binding properties of Cal0409 HA by analyzing its contacts with the α2-3 and α2-6 sialylated glycans.

On the basis of the observed contacts in the HA-glycan complexes, we summarize in **Table 1** the proposed roles of the residues in Cal0409 HA that provide binding specificity to α2-3 and α2-6 oligosaccharides, respectively. The main differences between the glycan-binding pockets of reference HAs and Cal0409 HA lie in the 140-loop region and the loop region preceding the 190-helix. Lys145, which is unique to Cal0409, along with Lys133 and Lys222, form a positively charged 'lysine fence' at the base of the binding site that potentially are positioned to anchor the N-acetylneuraminic acid (Neu5Ac) and galactose (Gal) sugars of both α2-3 and α2-6 glycans. In the case of the Cal0409 α2-6 oligosaccharide structural complex, the lysine fence also includes Lys156, which is positioned to provide additional contact with the glycan. The orientation of Asp190 is typically stabilized by a network of interactions involving residues at 186, 187 and 189 that precede the 190-helix. In Cal0409 HA, the residues at these positions are Ser186, Thr187 and Ala189; this set of residues is unique to the 2009 H1N1 strains. These residues appear to retain the ability to stabilize the orientation of Asp190 such that it is positioned to make optimal contacts with the third N-acetylglucosamine (GlcNAc) sugar (starting from Neu5Ac toward the reducing end) of α2-6 glycans, defined previously[8].

Our observations of the Cal0409 HA–glycan interactions suggest that this HA has the necessary residues to provide optimal contacts for

high affinity binding to α2-6 glycans present in the human upper airways[10]. Typically, the Glu190Asp substitution between avian and human-adapted H1 HA results in the loss of a critical contact with the Neu5Acα2-3Gal motif and a gain in contact with α2-6 glycans[1,5,8]. In the Cal0409 HA, however, this loss in contact to α2-3 glycans appears to be compensated for by Lys145 of the lysine fence. In summary, our analysis suggests that Cal0409 HA possesses residues that can be positioned to make optimal contacts with α2-6 (a characteristic binding feature shared by human H1N1 HAs[1,5,8]) as well as α2-3 sialylated glycans. In future studies, it will be important to experimentally determine the relative α2-6 and α2-3 binding affinities of the 2009 H1N1 HAs using appropriate methods.

Comparison of the antigenic regions[11] of Cal0409 NA with the consensus sequences of avian, human and swine-adapted N1 NAs shows that four positions—188, 331, 372 and 402—are novel in the 2009 H1N1 NA (**Supplementary Table 3** online). NA is presently the primary target of therapeutic intervention for influenza infection, and oseltamivir (Tamiflu) and zanamivir (Relenza) are widely used NA-inhibiting drug molecules. Recently, it has been reported that there is an alarming increase in the oseltamivir resistance of H1N1 viruses from 12.3% in 2007–2008 to 98.5% in 2008–2009 season (prior to the outbreak of the 2009 H1N1 infections)[12]. Fortunately, the 2009 H1N1 strains reported thus far are sensitive to both oseltamivir and zanamivir.

As the His274Tyr mutation is known to be responsible for resistance of the recent H1N1 human viruses to oseltamivir, we analyzed the potential effect of this mutation, should it occur, on the drug sensitivity of Cal0409 NA (**Fig. 2**). Previous studies have provided high-resolution crystal structures of NA–drug complexes, thereby shedding light on the structure-function relationships mediating the emergence of drug-resistant influenza strains[13]. Consistent with these studies, our analysis suggests that the bulky Tyr274 residue in the His274Tyr mutant form of Cal0409 NA is likely to displace Glu276 toward the docking site for the drug agent. The polar glycerol C6 group in zanamivir appears better suited than the corresponding hydrophobic pentyloxy group in oseltamivir to engage the resulting active site hydrogen bond network and thereby provides more optimal contacts with the His274Tyr mutant form of Cal0409 NA. The predicted unfavorable interaction of oseltamivir with the His274Tyr mutant form of Cal0409 NA is consistent with observations concerning the effect of this mutation in seasonal H1N1 strains and highlights the possibility of the emergence of Tamiflu-resistant viruses. Given the alarming proportion of circulating Tamiflu-resistant seasonal H1N1 viruses[13], our analysis support for both augmentation of oseltamivir stockpiles with additional drugs (including zanamivir) and the prudent administration of antivirals in general.

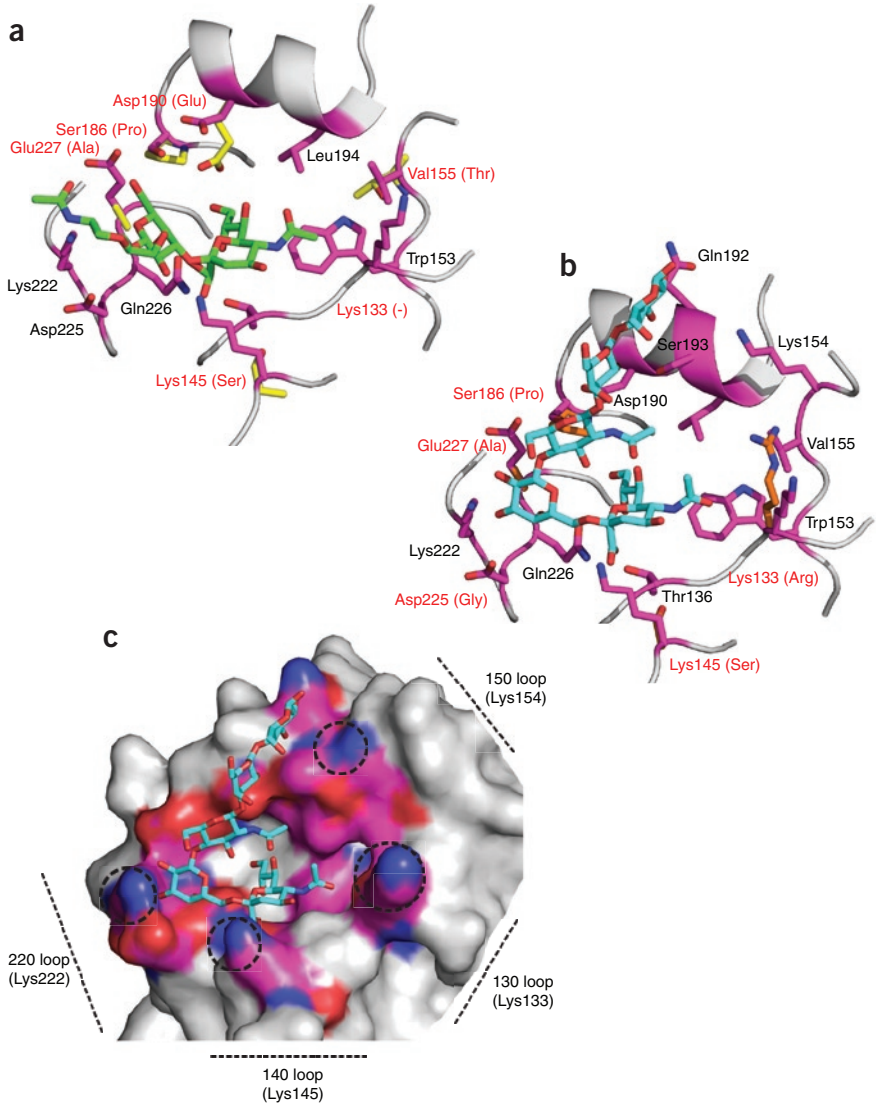


Figure 1 Glycan-binding properties of Cal0409 HA. A homology-based structural model of Cal0409 HA was constructed using the prototypic 1918 H1N1 HA (PDB ID: 1RUZ) as a template. (**a**) Theoretical structural model of Cal0409 HA (gray) bound to an α2-3 oligosaccharide (carbon; green) showing the key amino acids on HA (carbon; purple) that are positioned to make optimal contacts with the glycan. This complex was constructed using the PR8 (A/Puerto Rico/8/34) H1 HA–α2-3 oligosaccharide co-crystal structure (PDB ID: 1RVX; coordinates of trisaccharide Neu5Acα2-3Gal β1-4GlcNAc are ordered). The analogous amino acids in PR8 that are different from Cal0408 HA (labeled red with PR8 residues in parenthesis) are also shown (carbon; yellow). (**b**) Theoretical structural model of Cal0409 HA bound to an α2-6 oligosaccharide (carbon; cyan) showing the key amino acids on HA (carbon; purple) that are positioned to make optimal contacts with the glycan. This model was constructed using A/Swine/Iowa/30 (ASI30) H1 HA–α2-6 co-crystal structure (PDB ID: 1RVT; coordinates of pentasaccharide Neu5Acα2-6Galβ1-4GlcNAcβ1-3Gal β1-4Glc are ordered). The analogous amino acids in ASI30 HA that are different from Cal0408 HA (labeled red with ASI30 residues in parenthesis) are also shown (carbon; orange). The oxygen and nitrogen atoms are colored red and blue respectively. (**c**) Molecular surface of the HA-α-2-6 oligosaccharide complex highlighting the lysine fence (circled).

To probe the resistance of the 2009 H1N1 viruses to the adamantane-derivative drugs amantadine (Symmetrel) and rimantadine (Flumadine), we also investigated their target, the M2 protein. The prototypic mutation associated with adamantane resistance is Ser31Asn[14,15] and the presence of this mutation in all the 2009 H1N1 strains is consistent with their observed resistance to adamantane-derivative drugs[16].

To provide additional functional context to the above analysis, the amino acid sequences of

Table 1 Contacts observed in the structural model of Cal0409 HA–glycan complexes

Amino acid	α2-3	α2-6
Amino acids at base of HA pocket		
Tyr95, Thr136, Trp153, Val155, His183, Leu194, Gln226	Highly conserved residues in all H1 HAs involved in anchoring Neu5Ac	Highly conserved residues in all H1 HAs involved in anchoring Neu5Ac
Lys131, Lys145, Lys222	Lysine fence that makes optimal contacts with Neu5Ac α2-3Gal- motif	Lysine fence that makes optimal contacts with Neu5Acα2-6Gal- motif
Asp225	Minimal contact with α2-3	Optimal contact with Gal sugar of Neu5Acα2-6Gal motif
Glu227	Stabilize orientation of Lys222 side chain of the lysine fence	Stabilize orientation of Lys222 side chain of the lysine fence
Amino acids contacting additional sugars in α2-6 oligosaccharide		
Asp190	No contacts with α2-3	Optimal contact with the third GlcNAc sugar at reducing end of Neu5Acα2-6Gal- motif
Gln192, Ser193	No contacts with α2-3	Optimal contacts with sugars beyond third GlcNAc (toward reducing end)
Lys156	No contacts with α2-3	Part of the lysine fence positioned to make optimal contacts with sugars beyond the third GlcNAc (toward reducing end)
Ser186, Thr187, Ala189	No contacts with α2-3	Side chains of these residues form a network to stabilize orientation of Asp190

the HA and NA of the 2009 H1N1 strain were compared to those of the current vaccine H1N1 strain, A/Brisbane/59/07. A/Brisbane/59/07 has been associated with severe infectivity due, at least in part, to the fact that its HA and NA have low antigenic cross-reactivity to preexisting humoral immunity[17]. We then compared our results to the same analysis using A/Brisbane/59/07 and the previous H1N1 vaccine strain A/Solomon Islands/3/2006. This exercise provides an important reference point for a virus that is characterized as antigenically dissimilar to most circulating H1N1s. For the comparison between the NA of the 2009 H1N1 strain and the A/Brisbane/59/07 strain, we find that the overall percent sequence identity is 80.6%; however, this sequence identity drops to 38.0% within the antigenic regions (**Supplementary Table 4** online). A similar trend is observed in HA (**Supplementary Table 5** online), where the identity is reduced from 79.2% (overall) to 56.3% (antigenic). In contrast, when comparing A/Brisbane/59/07 with A/Solomon Islands/3/2006, the percent identity in NA shows only a marginal drop from 94.4% (overall) to 92.0% (antigenic). Similarly, with regards to HA alone, the percent identity

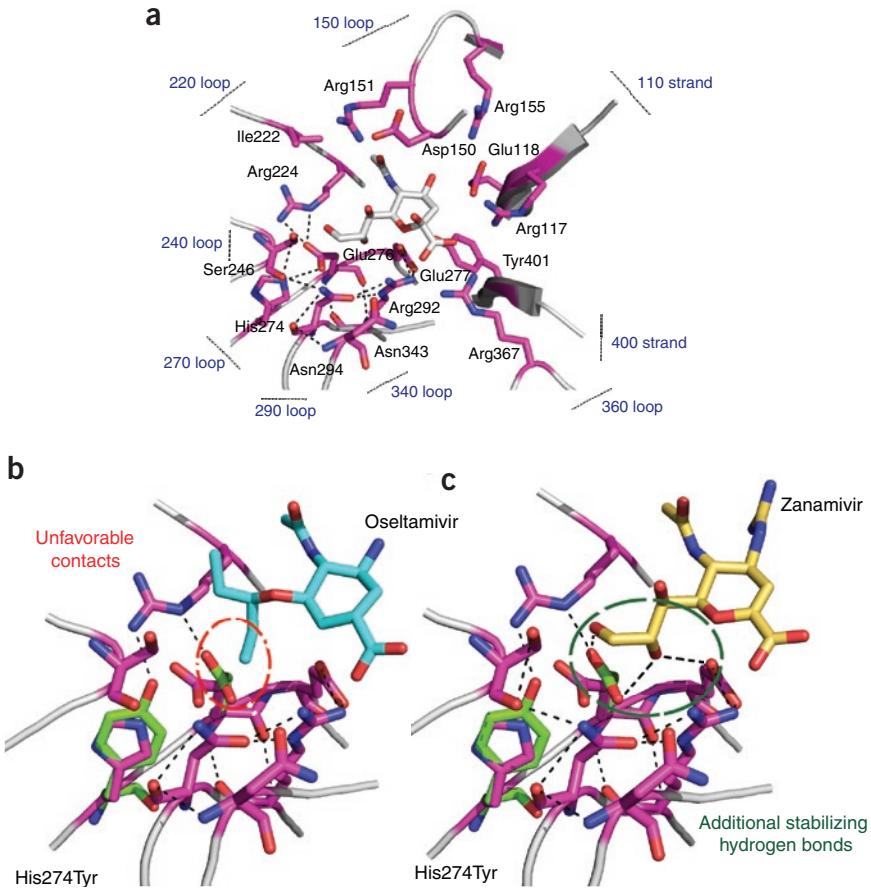


Figure 2 Analysis of Cal0409 NA. A homology-based structural model of Cal0409 NA was constructed using the N1 NA crystal structure (PDB ID: 3CKZ) as a template. (**a**) Active site of the Cal0409 neuraminidase (dark gray) docked with sialic acid (carbon; light gray) is shown highlighting key contacts including a dense network of hydrogen bonds (black broken lines) in the proximity of the C6 substitution (hydrogen atoms not displayed). Amino acid side chains are shown with carbon atoms colored purple. Oseltamivir and zanamivir were docked from their native co-crystal structures N1 NA–oseltamivir (PDB ID: 2HU0/2HU4) and N8 NA–zanamivir (PDB ID: 2HTQ) onto the Cal0409 NA structural model. (**b**) Comparison of contacts made by oseltamivir (carbon; cyan) with the wild-type (carbon; purple) and His274Tyr mutant (carbon; green) of Cal0409 NA shows the shift in Glu276 (caused by Tyr274 in the mutant) toward oseltamivir potentially resulting in unfavorable interactions (red broken circle) with the hydrophobic substitution in the C6 position. (**c**) Comparison of contacts made by zanamivir (carbon; yellow) with wild-type (carbon; purple) and His274Tyr mutant (carbon; green) of Cal0409 NA shows that the shift in the position of Glu276 toward the polar C6 substitution of zanamivir may provide additional favorable contacts (dark green broken circle). The oxygen and nitrogen atoms are colored red and blue, respectively, in all the structures.

drops from 98.6% (overall) to 97.2% (antigenic) (**Supplementary Tables 6** and 7 online). Taken together, the results indicate that the substantial variability observed in the antigenic regions of the 2009 H1N1 viruses would most likely result in the presentation of new epitopes that may not cross-react with the antibodies generated using the current vaccine strains, thereby potentially having important implications toward the protective effect afforded by existing seasonal influenza vaccines.

In addition to the vital role of the viral coat proteins, most prominently HA, in governing transmission, virulence and human adaptation, recent studies have demonstrated the critical role of the viral RNA polymerase PB2 in the efficient respiratory droplet (or airborne) transmission of wild-type human H1N1 viruses and avian-human reassorted influenza viruses in the ferret model[9]. In this study, it was demonstrated that a specific residue, Lys627, in human-adapted PB2 was critical for conferring efficient transmissibility. Conversely, mutation of this Lys627 to glutamic acid (which is typically found in avian and swine-adapted PB2) in PB2 of the 1918 pandemic strain (SC18) severely reduced its ability to transmit. Analysis of PB2 in all the 2009 H1N1 strains indicates that it has glutamic acid at position 627. On the basis of these earlier studies, we expect that the 2009 H1N1 viruses may be capable of transmission between humans, but the efficiency of transmission might be hampered by the absence of Lys627 in PB2. Although the HA of the 2009 H1N1 viruses is human-adapted, our analysis suggests that PB2 still requires an additional mutation to become fully human-adapted for efficient transmission.

A recent study[18] evaluated the pandemic potential of the 2009 H1N1 viruses using the available epidemiological data. This study concluded that despite the substantial uncertainty in the data, the clinical severity of the 2009 H1N1 viruses is more comparable to the 1957 H2N2 pandemic outbreak than the 1918 H1N1 pandemic. Given the fact that the evolution of this virus is uncertain at best, we must remain vigilant for additional mutations that can render this strain more virulent.

In conclusion, we set out using sequence information to evaluate the HA, NA PB2 and M2 genes in the new H1N1 viral strain based on their known critical roles in the human adaptation, human-to-human transmission and resistance to currently used antiviral drugs (**Table 2**). As experimental data for 2009 H1N1 viruses become available, it will be possible to correlate those results with the analyses presented here.

Table 2 Summary of predicted properties of HA, NA, PB2 and M2 for current 2009 H1N1 viruses

Protein	Predicted properties	Perspectives
Hemagglutinin (HA)	• Glycan-binding site (including Asp190 and Asp225) is similar to that of human adapted H1N1 HAs • Binding site contains residues that are positioned to make optimal contacts with both α2-6 and α2-3 glycans • Lys145, Ser186, Thr187 and Ala189 are novel substitutions that have not been observed in other human H1N1 HAs	• HA is human adapted and is expected to bind with high affinity to α2-6 and better affinity to α2-3 in comparison with other human H1N1 HAs • Novel substitutions in the antigenic regions of HA might present new epitopes
Neuraminidase (NA)	• Novel substitutions in four positions 189, 331, 369 and 398 in the putative antigenic site • Active site has not yet acquired the characteristic mutations such as His274Tyr that provides resistance to oseltamivir. • Zanamivir can potentially make optimal contacts with the oseltamivir-resistant NA mutants.	• Novel substitutions in antigenic regions of NA might present new epitopes • NA can acquire mutations that offer resistance to Tamiflu • Zanamivir might be preferred to oseltamivir
PB2 and M2	• Lys627Glu mutation in PB2 reduced transmission efficiency of a prototypic human-adapted H1N1 virus[9] • M2 protein has acquired Ser31Asn mutation that provides resistance to adamantane-derived drugs	• Acquisition of the Glu627Lys mutation could potentially improve transmission efficiency of the virus

Note: Supplementary information is available on the Nature Biotechnology website.

ACKNOWLEDGMENTS

The authors would like to acknowledge support from the Singapore–Massachusetts Institute of Technology Alliance for Research and Technology; and the National Institute of General Medical Sciences of the National Institutes of Health (GM 57073 andU54 GM62116) to R.S.

***Venkataramanan Soundararajan*[1,2]*, Kannan Tharakaraman*[1,2]*, Rahul Raman*[1]*, S Raguram*[1]*, Zachary Shriver*[1]*, V Sasisekharan*[1] *& Ram Sasisekharan*[1]

[1]*Harvard-MIT Division of Health Sciences & Technology, Koch Institute for Integrative Cancer Research and Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.*

[2]*These authors contributed equally to this work.*
e-mail: rams@mit.edu

1. Gamblin, S.J. *et al. Science* **303**, 1838–1842 (2004).
2. Taubenberger, J.K. *et al. Nature* **437**, 889–893 (2005).
3. Tumpey, T.M. *et al. Science* **310**, 77–80 (2005).
4. Palese, P., Tumpey, T.M. & Garcia-Sastre, A. *Immunity* **24**, 121–124 (2006).
5. Stevens, J. *et al. J. Mol. Biol.* **355**, 1143–1155 (2006).
6. Tumpey, T.M. *et al. Science* **315**, 655–659 (2007).
7. Pappas, C. *et al. Proc. Natl. Acad. Sci. USA* **105**, 3064–3069 (2008).
8. Srinivasan, A. *et al. Proc. Natl. Acad. Sci. USA* **105**, 2800–2805 (2008).
9. Van Hoeven, N. *et al. Proc. Natl. Acad. Sci. USA* **106**, 3366–3371 (2009).
10. Chandrasekaran, A. *et al. Nat. Biotechnol.* **26**, 107–113 (2008).
11. Fanning, T.G., Reid, A.H. & Taubenberger, J.K. *Virology* **276**, 417–423 (2000).
12. Dharan, N.J. *et al. J. Am. Med. Assoc.* **301**, 1034–1041 (2009).
13. Collins, P.J. *et al. Nature* **453**, 1258–1261 (2008).
14. Stouffer, A.L. *et al. Nature* **451**, 596–599 (2008).
15. Pielak, R.M., Schnell, J.R. & Chou, J.J. *Proc. Natl. Acad. Sci. USA* **106**, 7379–7384 (2009).
16. Gubareva, L. *et al. Morb. Mortal Wkly. Rep.* **58**, 433–435 (2009).
17. Fiore, A.E. *et al. MMWR Recomm. Rep.* **57**, 1–60 (2008).
18. Fraser, C. *et al. Science* DOI: 10.1126/science.1176062, May 14, 2009.

Science communication reconsidered

Tania Bubela[1,2*], Matthew C Nisbet[3], Rick Borchelt[4], Fern Brunger[5], Cristine Critchley[6], Edna Einsiedel[7], Gail Geller[8–10], Anil Gupta[11], Jürgen Hampel[12], Robyn Hyde-Lay[2,13], Eric W Jandciu[14], S Ashley Jones[15], Pam Kolopack[16], Summer Lane[2], Tim Lougheed[17], Brigitte Nerlich[18], Ubaka Ogbogu[2,19], Kathleen O'Riordan[20,21], Colin Ouellette[2], Mike Spear[13], Stephen Strauss[22], Thushaanthini Thavaratnam[23], Lisa Willemse[24] & Timothy Caulfield[25–27]

As new media proliferate and the public's trust and engagement in science are influenced by industry involvement in academic research, an interdisciplinary workshop provides some recommendations to enhance science communication.

Science communication receives significant attention from policy makers, research institutions, practitioners and scholars[1,2]. It is a complex and contentious topic that encompasses a spectrum of issues from the factual dissemination of scientific research to new models of public engagement whereby lay persons are encouraged to participate in science debates and policy.

Over the past several decades, the complexities of science communication have been magnified by institutional, social and technological change. Science increasingly is interdisciplinary, bureaucratic, global in scale, problem-based and dependent on private funding. This latter trend, in particular, raises issues of public trust in science, which studies have shown is diminished by researcher and institutional affiliation with the private sector, especially in the area of biomedicine[3,4].

Technology has also transformed the nature of the media system, creating an abundance of cable television, Internet and digital resources for the public to inform themselves about science and its social implications. With these new outlets, highly motivated individuals have a greater ability to learn about science and to become involved in collective decision-making[5]. Yet media fragmentation also means that if individuals lack an interest in science, they can very easily avoid science media altogether. There is a general concern that reduced quality of reporting by some media sources, primarily television and online, may have negative impacts, such as demands for inappropriately hyped medical services[6,7].

Jean-Francois Podevin/Science Photo Library

Science communication faces stiff challenges with the blurring of boundaries between public and private science and the fragmentation of audiences.

With this convergence of social forces and journalistic challenges in mind, we convened an interdisciplinary workshop on the changing nature of science communication, focusing specifically on biotech, biomedicine and genetics. What follows is a discussion of the questions and issues addressed by experts from the US, the UK, Canada, Germany and Australia. Our goal is to focus attention on key areas of expert agreement about two aspects of science communication: public engagement and science journalism. These two main themes are interrelated; the dissemination of knowledge is one part of a multifaceted approach toward increasing public involvement in science issues and decision-making. We conclude with specific recommendations for moving forward.

Models and assumptions guiding science communication

Despite increasing attention to new directions in public engagement, a still-dominant

*A list of affiliations appears at the end of the paper.
e-mail: nisbet@american.edu

assumption among many scientists and policy-makers is that when controversies over science occur, ignorance is at the root of public opposition. Concerns are raised about the state of science education and scientific literacy more generally[8,9]. Science communication initiatives are therefore directed at filling in the 'deficit' in knowledge, with the hope that if members of the public only understood the scientific facts, they would be more likely to see the issues as experts do. The strategy is thus to inform the public by way of popular science outlets such as television documentaries, science magazines, newspaper science coverage and more recently science websites and blogs.

Of course, some knowledge about science, and especially its role in society, is fundamentally important for a public that bears the risks and benefits of scientific and technological development[10]. Yet the narrow emphasis of the deficit approach does not recognize that knowledge is only one factor among many influences that are likely to guide how individuals reach judgments, with ideology, social identity and trust often having stronger impacts[10]. The deficit model also overlooks the fact that, given the abundance of competing content choices, traditional science media outlets reach only a relatively small audience of already knowledgeable science enthusiasts. In addition, on certain topics, such as cloning, the public is likely to draw strongly upon the portrayals featured in entertainment film and television, science fiction novels and other forms of popular culture[11–13].

A decade ago, a new 'public engagement' or interactive model emerged—one that emphasizes deliberative contexts in which a variety of stakeholders can participate in a dialog so that a plurality of views can inform research priorities and science policy[1]. These efforts toward two-way dialog with lay publics have taken various forms, such as deliberative polls, citizen juries, consensus conferences and *cafés scientifiques*. As a participatory process, each form might place a different weight on 'extended peer review,' whereby the 'publics,' or groups of individuals who are affected by the products of science, are invited to become part of a community of evaluators and decision-makers. Initiatives also vary in terms of how participants are asked for feedback, how much their feedback influences the final decisions and the timing of consultation[14].

Studies find that lay participants not only learn directly about the technical aspects of a subject, such as food biotech or biomedical research, but also learn about the social, ethical and economic implications of the science. Participants also feel more confident and efficacious in their ability to participate in science decisions, perceive scientists and their organizations as more responsive to their concerns, and say afterwards that they are motivated to become active on the issue if provided a future opportunity to do so[15,16].

Advocates for expanding these public engagement initiatives argue that consultation exercises often come too late (usually just as a science product, such as nanotechnology, is being introduced to the market), that lay input is not given enough weight in decision-making and that under these conditions the consultation process only serves a public relations function. They argue that engagement needs to move 'upstream' to when science or technology is in its formative stage, so that relevant publics can have a more meaningful say in matters of ownership, regulation, uses, benefits and risks[17–19]. Given this, the media could play an important role in informing the public about early-stage science policy debates and avenues for public involvement, potentially raising awareness and participation[20]. Yet a genuine role for lay participants' recommendations can come only with the realization that sometimes an engaged public might reach collective decisions that go against the self-interests of scientists. For example, one outcome of a recent consultation forum on nanotechnology was that several lay participants were motivated to form an advocacy group to act as a watchdog over research in their community[15].

Framing the message

The deficit model blames failures in science communication on inaccuracies in news coverage and the irrational beliefs of the public, but it ignores several realities about audiences and how they use the media to make sense of science. First, individuals are naturally 'cognitive misers': if they lack a motivation to pay close attention to science debates, they will rely heavily on mental shortcuts, values and emotions to make sense of an issue, often in the absence of knowledge[21,22]. Second, as part of this miserly nature, individuals are drawn to news sources that confirm and reinforce their pre-existing beliefs. This tendency, of course, has been facilitated by the fragmentation of the media and the rise of ideologically slanted news outlets[23]. Third, opinion leaders other than scientists, such as religious leaders, nongovernmental organizations and politicians, have been successful in formulating their messages about science in a manner that connects with key stakeholders and publics but at times might directly contradict scientific consensus or cut against the interests of organized science[24].

Under these conditions, audiences will pay more attention to certain dimensions of a science debate over others depending on how an issue is 'framed' in news coverage. Frames are interpretative packages and storylines that help communicate why an issue might be a problem, who or what might be responsible and what should be done[25]. Frames are used by lay publics as interpretative schemas to make sense of and discuss an issue; by journalists to condense complex events into interesting and appealing news reports; by policy-makers to define policy options and reach decisions; and by scientists to communicate the relevance of their findings. In each of these contexts, frames simplify complex issues by lending greater weight to certain considerations and arguments over others[26]. Framing is an unavoidable reality of the science communication process.

There is growing awareness among science organizations that if they want to be more effective at using the media to communicate with a diversity of audiences, they need to switch the frame—or interpretative lens—by which they communicate about a scientific topic, such as evolution, stem cell research or nanotechnology[27]. Instead of relying on personal experience or anecdotal observation, it is necessary to carry out careful audience research to determine which frames work across intended audiences. Communication is both an art and a science. For example, the US National Academies (Washington, DC) used focus groups and polling to inform the structure of a written report about the teaching of evolution and to plan publicity efforts. Their research indicated that an effective storyline for translating the relevance of evolutionary science for students was one emphasizing the connection to advances in modern medicine. Contrary to their expectations, the research concluded that an alternative frame emphasizing recent court decisions did not provide nearly as effective a message[28].

Yet turning to audience research requires a delicate balance on the part of science organizations. Any reframing of an issue needs to remain true to the state of the underlying science. For example, in promoting human embryonic stem cell research around the 'hope for cures', some advocates have given the false impression that available therapies are just a few years away, an interpretation that puts public trust at risk. Similarly, some industry advocates have re-framed food biotech as a moral quest to improve global food security, but their promise of 'putting an end to world hunger' dramatically oversimplifies a complex problem[29].

The challenges of science journalism

The media not only influence public perceptions but also shape and reflect the policy debate[30]. Few decisions are made by

policymakers and stakeholders without the media in mind. Given this role and influence, there have long been concerns about distortion and hype in news coverage of biomedicine and biotech. The orientation toward hype is viewed internationally by many scientists, ethicists, policymakers and government officials as the primary shortcoming of the media.

In general, there is a stable baseline level of media coverage of biomedicine and biotech. Much of this news attention is driven by a small number of prestigious and highly influential scientific journals, with science framed in this coverage in terms of social progress and economic growth[31–33]. Numerous studies of media content have shown that coverage in newspapers is surprisingly accurate, with few errors of commission[31,34]. Assessing accuracy in the reporting of a single study, however, does not address whether the coverage contextualizes where the study fits within an emerging body of knowledge, drawing comparisons to other studies or expert views. Thus, as a caveat, accuracy in reporting and the dissemination of high-quality evidence are not necessarily synonymous[33].

In regard to perceptions of coverage, contrary to conventional wisdom, research has consistently shown that most scientists are satisfied with the media coverage of their own research and are more likely to be critical of science coverage generally[35]. Research similarly suggests that perceptions of bias in the coverage of biotech vary depending on a stakeholder's connection and personal commitment to the topic[36].

Studies have shown that hype in the media is most likely to originate with researchers using metaphors associated with breakthroughs[37] when in reality their research is one more incremental piece of a complex scientific endeavor. Prominent scientists certainly contribute to the creation of overly positive or negative expectations[38]. Numerous commentators have remarked that the media, scientists, the public and other interest groups can become complicit in generating a 'cycle of hype'[39]. The cycle is driven by enthusiastic researchers facing pressures from their research institutions, funders and industry; by the desire of institutions and journals to bolster their profiles; by a profit-driven media; and by the need of individual journalists to define events as newsworthy[39,40]. As one result of these factors, research has shown that positive results are more likely to be published[41], whereas studies that refute previously published research are less likely to gain attention. For example, the discovery of the 'gay gene' was published in *Nature* and received considerable media attention[42,43], but a study refuting these findings received limited press coverage[43].

A further source of hype may lie in errors of omission—what is left out of media narratives[34,44]. There is a lack of reporting on funding sources for research and potential conflicts of interest, information essential for the lay public to assess the credibility of the research[45,46] and which group of experts to trust. Public opinion surveys indicate a high degree of trust in scientists generally and university scientists specifically, but this trust declines when members of the public are asked their impression of industry scientists[3]. Comfort with a technology increases with public trust in regulatory authorities and government. In fact, unless a science issue is contested by rival cultural authorities, such as religious or political leaders, the public tends to defer strongly to the expertise of university and government scientists[47].

Details of methods and study design (especially for clinical trials), risks and timelines for the delivery of benefits are also underreported. Risks are often underreported because of the difficulties of conveying probabilistic information, which is inadequately understood by most journalists and by the general public[31,34]. However, it is not just probabilistic risks that are underplayed but also any broader discussion of social and ethical risks of the research. Equally of concern is the lack of discussion about realistic timelines for the delivery of benefits arising from what, in most cases, is still early-stage research. Omitting timelines may produce an impression in the public's mind that significant therapeutic benefits are imminent—the lay public and experts have very different perceptions of timelines. This is particularly dangerous in regard to stem cell research where people are desperate to gain access to stem cell therapies or 'miracle cures'.

The caveat about these previous content analysis studies is that the majority have concentrated on the print media, and primarily just the science beat, ignoring the fact that the media are not homogeneous. This approach ignores the degree to which local and national television news broadcasts, and increasingly the Internet, are now primary sources of public affairs information for the public[48]. Studies have also tended to focus narrowly on science journalists, but science debates receive their greatest attention when they shift from being covered just by these specialists to become the focus of political journalists, commentators and pundits. Under these conditions, the image of science morphs from a focus on discoveries packaged as progress, promise and technical background to a new emphasis on conflict and dramatic claims about risks and ethics[29,49].

This difference in perception, and the hype derived from errors of omission and framing, may already be leading to individual and social harm. The public has access to commercially available genetic tests marketed directly to consumers, which provide health information in the form of probabilistic risk factors[50,51], and to as-yet-unapproved stem cell therapies in jurisdictions with lower regulatory standards[52]. This raises important questions about the roles and responsibilities of the media.

Media roles and responsibilities

Many academic articles, editorials and reports draw on findings about errors of omission and accuracy to recommend best practices and checklists for journalists[53–55]. But do such endeavors confront the realities of science journalism and other news beats? The most important issue may not necessarily be content, but rather how the research is framed. In this regard, it is critical to understand the factors that shape the dominant interpretations in news coverage.

First, there is often a fundamental disconnect between how scientists and journalists interpret and describe the research process. For example, scientific papers are relentlessly quantitative, whereas media articles are often based on humanized accounts designed to connect with lay readers. Scientific articles are aimed at a narrow specialist audience, whereas media articles are aimed at a broader audience. As a result, journalistic accounts are based on personal anecdotes provided by researchers or by individuals who may directly benefit from the research, such as affected individuals or members of affected families. Without such connections, science stories are less likely to be published in competition with the news of the day.

New media are also fundamentally changing the nature of science communication. The role of the Internet as a major source of biomedical and science information for the public has both positive and negative consequences. Traditional media websites allow journalists to connect readers with source information through direct links to research or patient sites and articles. The expanded layout of web pages may address concerns about errors of omission, as more quantitative or probabilistic information may be provided in sidebars or graphics but only if the effort is made to provide this sometimes labor-intensive material. Special online comment sections allow readers to instantly contest or correct information contained in a story. Scientists and science journalists who double as bloggers provide readers with background and context about specialized areas of research. Science blogs create a dialog with readers, merging online interaction with real-world socializing at *cafés scientifiques*

and other informal settings. Science bloggers frequently vet false claims made in the media or in policy debates and increasingly serve as important sources for journalists.

However, much of the information on the Internet comes from sources other than the mainstream media or scientist bloggers, and much of this may be of dubious quality. Corporate information sources generally are little more than direct-to-consumer advertising for products, services or both. For example, nutrigenomic testing services offered on the Internet are often tied to the sale of nutriceuticals and other products[56,57]. Only recently have corporations begun to take advantage of the social media properties of the web, entering into a dialog with stakeholders and publics via specially created sites that feature blogs, scientist profiles and discussion sections (see Johnson & Johnson's (Bridgewater, NJ, USA) corporate blog (http://jnjbtw.com), YouTube channel (http://www.youtube.com/user/JNJhealth) and Facebook page (http://www.facebook.com/ADHDMoms)). Other sites cater to special interest groups—for example, creationist or anti–stem cell research websites on the one hand and atheist or patient advocacy groups on the other—and are intended to strategically frame news coverage and/or the policy debate. Science blogs also engage in strategic framing, with some of the most popular science bloggers blending discussion of science with ideologically driven commentary on politics or religion. These popular blog sites become echo chambers reinforcing deficit-model assumptions about the public, singling out science literacy as the golden key to winning public support and to eroding religious belief.

Finally, the greatest challenge to science communication online remains simply *reaching audiences*. The availability of science information from credible sources online does not mean the public will use it. Even more than with the traditional media, if people lack an interest in science content on the web, they can very easily ignore it. This has implications for the public's degree of engagement with science policy debates.

Recommendations and challenges

The proliferation of information sources combined with increased industrial involvement in scientific research raise the issue of public trust and engagement with science. The primary concerns are the blurring of boundaries between public and private science and the fragmentation of audiences. Science communication, therefore, remains driven by an ever-more-complex relationship between institutions, stakeholders, the media and a diversity of publics.

In this context, clarification about the goals and assumptions of science communication is required, recognizing the complexity and variety of issues to be communicated. Current initiatives toward public education and involvement are presented as representing democratic reforms and being more inclusionary than past efforts, yet remain based on the deficit model, which research has shown to be insufficient. On this matter, then, there needs to be continued investment in public dialog initiatives, such as deliberative forums and consensus conferences. Yet, importantly, the focus of these deliberative exercises should be an honest effort at relationship- and trust-building[58] rather than persuasion, with mechanisms for actively incorporating the input of lay participants into decision-making[59].

When it comes to effectively working with media organizations to engage key audiences, it is necessary to recognize the importance of framing as well as the differing assumptions and imperatives of scientists, journalists and key publics. Public trust and the perception of media portrayals will vary by an individual's social identity and values. Science communication efforts should therefore be supported by careful audience research, such as that done by the National Academies on evolution. This strategy does not mean engaging in false spin or hype, but rather involves drawing upon research to explore alternative storylines, metaphors and examples that more effectively communicate both the nature and the relevance of a scientific topic, such as human embryonic stem cell research.

Graduate students, as the future spokespeople and decision-makers at science institutions, should be taught about the social and political context of science and how to communicate with the media and a diversity of publics. The latter includes an emphasis on the importance of meaningful public dialog initiatives as well as of relationship-building with journalists and editors[60]. There is a danger, however, of this type of public engagement emphasis becoming too conflated with marketing and public relations.

The wide-ranging factors contributing to media hype and errors (largely of omission) need to be more explicitly recognized so as to allow science institutions and media organizations to formulate appropriately informed communication policies.

To enhance our understanding of science communication in the context of new media, the focus of research on science communication should be expanded to include online and digital media, while recognizing the continued agenda-setting nature of traditional news sources. Given the fragmented nature of Internet audiences, if organizations want to broaden their reach when producing science content online, they need to find ways to facilitate incidental exposure, gaining the attention of key publics at places on the web where they are not actively looking for science information. There also will need to be laws protecting consumers from false or hyped claims on websites that market health services and products directly to the public.

Much as we have ever-improving measures of public opinion about science and an increasing number of survey data sources and studies to reference, there also needs to be investment in the systematic tracking of news and cultural indicators, including traditional news outlets but also talk radio, late-night satirical programming, religious media, the web and new documentary genres as well as entertainment television and film. Each of these media zones may constitute a different cultural context in which the public will interpret science.

At journalism schools and news organizations, the development of a new 'science policy' beat should be encouraged. This will fill in the gaps between the technical backgrounders preferred by science writers and the conflict emphasis of political reporters, providing important background for debates on science policy. In this context, discussion of science as a social institution could include funding structures, public-private institutional relationships and commercialization. An open public discussion of the blurring public-private divide in science could only enhance public trust.

Finally, if there is a major threat to science journalism, it is that science journalists are losing their jobs at for-profit news organizations. Some suggest that scientists-as-bloggers might be able to fill the gap[61], yet for reasons reviewed earlier, this is unlikely to be an effective solution. New models of foundation-, university- or government-supported science journalism are needed, with these online digital formats blending professional reporting with user-generated content and discussion.

ACKNOWLEDGMENTS

The workshop was supported by grants from Genome Canada (grant no. G53400000 to T. Caulfield, E. Einsiedel, P. Phillips and M. Veeman), The Stem Cell Network (to T. Caulfield, E. Einsiedel and B.M. Knoppers) and The Advanced Food and Materials Network (grant no. 18G to T. Caulfield and D. Castle). We thank N. Hawkins at the Health Law Institute (Edmonton, Alberta, Canada) for administrative support.

1. House of Lords. *Science and Society* (House of Lords, London, 2000). <http://www.parliament.the-stationery-office.co.uk/pa/ld199900/ldselect/ldsctech/38/3801.htm>.

2. The Royal Society. *Factors Affecting Science Communication: A Survey of Scientists and Engineers* (The Royal Society, London, 2006). <http://royalsociety.org/page.asp?id=3180 2>.
3. Critchley, C.R. *Public Underst. Sci.* **17**, 309–327 (2008).
4. United Kingdom Research Councils. *UK Public Attitudes to Science, 2008: A Survey* (RCUK, Swindon, 2008). <http://www.rcuk.ac.uk/cmsweb/downloads/rcuk/scisoc/pas08.pdf>.
5. <http://well.blogs.nytimes.com/2008/09/29/decoding-your-health/>.
6. Orkin, S.H. & Motulsky, A.G. *Report and Recommendations of the Panel to Assess the NIH Investment in Research on Gene Therapy* (National Institutes of Health, Bethesda, Maryland, USA, 1995). <http://www.nih.gov/news/panelrep.html>.
7. Stockdale, A. *Sociol. Health Illn.* **21**, 579–596 (1999).
8. National Science Foundation. *Science and Technology: Public Attitudes and Public Understanding* (National Science Board, Arlington, Virginia, USA, 1998).
9. UK Office of Science and Technology Science and the Public. *A Review of Science Communication and Attitudes to Science in Britain* (Wellcome Trust, London, 2000).
10. Sturgis, P. & Allum, N. *Public Underst. Sci.* **13**, 55–74 (2004).
11. Haran, J., Kitzinger, J., McNeil, M. & O'Riordan, K. *Human Cloning in the Media: From Science Fiction to Science Practice* (Routledge, Abingdon, UK, 2007).
12. Nisbet, M.C. & Goidel, K. *Public Underst. Sci.* **16**, 421–440 (2007).
13. Nerlich, B., Clarke, D.D. & Dingwall, R. *Soc. Res. Online* **4** (1999) <http://www.socresonline.org.uk/socresonline/4/3/Nerlich.htm>.
14. Einsiedel, E. Public engagement and dialogue: a research review. in *Handbook of Public Communication on Science and Technology* (eds. Bucchi, M. & Smart, B.) 173–184 (Routledge, London, 2008).
15. Powell, M. & Kleinman, D.L. *Public Underst. Sci.* **17**, 329–348 (2008).
16. Besley, J.C., Kramer, V.L., Yao, Q. & Tourney, C.P. *Sci. Commun.* **30**, 209–235 (2008).
17. Wynne, B. *Community Genet.* **9**, 211–220 (2006).
18. Wilsdon, J. & Willis, R. *See-through Science: Why Public Engagement Needs to Move Upstream* (Demos, London, 2004).
19. Rogers-Hayden, T. & Pidgion, N. *Public Underst. Sci.* **16**, 345–364 (2007).
20. Goidel, K. & Nisbet, M.C. *Polit. Behav.* **28**, 175–192 (2006).
21. Downs, A. *An Economic Theory of Democracy* (Harper, New York, 1957).
22. Popkin, S. *The Reasoning Voter* (University of Chicago Press, Chicago, 1991).
23. Mutz, D. in *Red and Blue Nation*, vol. 1 (eds. Nivola, P. & Brady, D.W.) 222–263 (The Brookings Institution, Washington, DC, 2006).
24. Nisbet, M.C. & Mooney, C. *Science* **316**, 56 (2007).
25. Gamson, W.A. & Modigliani, A. *Am. J. Sociol.* **95**, 1–37 (1989).
26. Scheufele, D.A.J. *Communication* **49**, 103–122 (1999).
27. Nisbet, M.C. & Scheufele, D.A. *Scientist* **21**, 39–44 (2007).
28. Labov, J.B. & Kline Pope, B. *CBE Life Sci. Educ.* **7**, 20–24 (2008).
29. Nisbet, M.C. & Huge, M. *Int. J. Press/Politics* **11**, 2, 3–40 (2006).
30. Caulfield, T., Bubela, T. & Murdoch, C. *Genet. Med.* **9**, 850–855 (2007).
31. Bubela, T. & Caulfield, T. *Can. Med. Assoc. J.* **170**, 1399–1407 (2004).
32. Nisbet, M.C. & Lewenstein, B.V. *Sci. Commun.* **23**, 359–391 (2002).
33. Durant, J., Bauer, M. & Gaskell, G. *Biotechnology in the Public Sphere: A European Sourcebook* (Michigan State University Press, Lansing, Michigan, USA, 1998).
34. Holtzman, N.A. *et al. Community Genet.* **8**, 133–144 (2005).
35. Peters, H.P. *et al. Sci. Commun.* **321**, 204–205 (2008).
36. Gunther, A.C. & Schmitt, K. *J. Commun.* **54**, 55–70 (2004).
37. Nerlich, B. in *Cognitive Foundations of Linguistic Usage Patterns* (eds. Schmid, H.J. & Handl, S.) (Mouton de Gruyter, Berlin, in the press).
38. Nerlich, B. & Halliday, C. *Sociol. Health Illn.* **29**, 46–65 (2007).
39. Caulfield, T. *Trends Biotechnol.* **22**, 337–339 (2004).
40. Bubela, T. *Clin. Genet.* **70**, 445–450 (2006).
41. Vickers, A., Goyal, N., Harland, R. & Rees, R. *Control. Clin. Trials* **19**, 159–166 (1998).
42. Conrad, P. & Markens, S. *Health* **5**, 373–390 (2001).
43. Petersen, A. *J. Commun. Inq.* **23**, 163–182 (1999).
44. Mountcastle-Shah, E. *et al. Sci. Commun.* **24**, 458–478 (2003).
45. Cook, D.M., Boyd, E.A., Grossmann, C. & Bero, L.A. *PLoS One* **2**, e1266 (2007).
46. McComas, K.A. & Simone, L.M. *Sci. Commun.* **24**, 395–419 (2003).
47. Brossard, D. & Nisbet, M.C. *Int. J. Public Opin. Res.* **19**, 24–52 (2007).
48. <http://pewresearch.org/pubs/928/key-news-audiences-now-blend-online-and-traditional-sources>
49. Jasanoff, S. *Nature* **450**, 33 (2007).
50. Gollust, S.E., Wilfond, B.S. & Hull, S.C. *Genet. Med.* **5**, 332–337 (2003).
51. Mayo Clinic Staff. Genetic testing you can order online. *Women's Health* (Mayo Foundation for Medical Education and Research, 26 March 2008). <http://www.mayoclinic.com/health/genetic-testing/GA00058>.
52. Lau, D. *et al. Cell Stem Cell* **3**, 591–594 (2008).
53. Blum, D., Knudson, M. & Marantz Henig, R. (eds.). *A Field Guide for Science Writers*. (Oxford University Press, Oxford, UK, 2005).
54. Schwitzer, G. *et al. PLoS Med.* **2**, e215 (2005).
55. Thompson, L. in *Genes and Human Self-Knowledge* (eds. Weir, R., Lawrence, S.C. & Fales, E.) 104–121 (University of Iowa Press, Iowa City, Iowa, USA, 1994).
56. Bubela, T. & Taylor, B. *Health Law Rev.* **16**, 39–47 (2008).
57. U.S. Government Accountability Office. *Nutrigenetic Testing: Tests Purchased from Four Web Sites Mislead Consumers* (GAO-06–977T, 2006). <http://www.gao.gov/products/GAO-06–977T>.
58. Yarborough, M., Fryer-Edwards, K., Geller, G. & Sharp, R.R. *Acad. Med.* (in the press).
59. Borchelt, R. & Hudson, K. *Sci. Prog.* Spring/Summer: 78–81 (2008).
60. Geller, G., Bernhardt, B.A., Rodgers, J.E. & Holtzman, N.A. *Genet. Med.* **7**, 198–205 (2005).
61. Brumfield, G. *Nature* **458**, 274–277 (2008).

[1]School of Public Health and [2]Health Law Institute, Law Centre, University of Alberta, Edmonton, Alberta, Canada. [3]School of Communication, American University, Washington, DC, USA. [4]Genetics and Public Policy Center, Johns Hopkins University, Washington, DC, USA. [5]Division of Community Health & Humanities, Faculty of Medicine, Memorial University of Newfoundland, St. John's, Newfoundland, Canada. [6]Faculty of Life and Social Sciences, Swinburne University of Technology, Hawthorn, Victoria, Australia. [7]Faculty of Communication and Culture, University of Calgary, Calgary, Alberta, Canada. [8]Berman Institute of Bioethics and [9]Department of Medicine, School of Medicine, Johns Hopkins University, Baltimore, Maryland, USA. [10]Department of Health, Behavior and Society, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA. [11]Applied Research and Analysis Directorate, Health Policy Branch, Health Canada, Ottawa, Ontario, Canada. [12]Department for Sociology of Technology and Environment, University of Stuttgart, Stuttgart, Germany. [13]Genome Alberta, Calgary, Alberta, Canada. [14]Science Journalism Research Group, School of Journalism, University of British Columbia, Vancouver, British Columbia, Canada. [15]Faculty of Communication and Culture, University of Calgary, Calgary, Alberta, Canada. [16]Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada. [17]Canadian Science Writers Association, Toronto, Ontario, Canada. [18]Institute for Science and Society, University of Nottingham, Nottingham, UK. [19]Faculty of Law, University of Toronto, Toronto, Ontario, Canada. [20]Centre for Material Digital Culture, University of Sussex, Brighton, UK. [21]ESRC Centre for Economic and Social Aspects of Genomics (Cesagen), Institute for Advanced Studies, County South, Lancaster University, Lancaster, UK. [22]Canadian Broadcasting Corporation, Toronto, Ontario, Canada. [23]Faculty of Communication and Culture, University of Calgary, Calgary, Alberta, Canada. [24]Stem Cell Network, Ottawa, Ontario, Canada. [25]Health Law Institute, [26]Faculty of Law and [27]School of Public Health, University of Alberta, Edmonton, Alberta, Canada.

Strangled at birth? Forest biotech and the Convention on Biological Diversity

Steven H Strauss, Huimin Tan, Wout Boerjan & Roger Sedjo

Against the Cartagena Protocol and widespread scientific support for a case-by-case approach to regulation, the Convention on Biological Diversity has become a platform for imposing broad restrictions on research and development of all types of transgenic trees.

The Convention on Biological Diversity (CBD) has become a major focus of activist groups that wish to ban field research and commercial development of all types of genetically modified (GM) trees. Recent efforts to influence CBD recommendations by such groups has led to the adoption of recommendations for increased regulatory stringency that are inconsistent with the views of most scientists and most of the major environmental organizations. We suggest that the increasingly stringent recommendations adopted by the CBD in recent years are impeding, and in many places may foreclose, much of the field research needed to develop useful and safe applications of transgenic trees. To move forward, improvements to regulations are needed that allow field research to be conducted at a reasonable cost and under workable levels of confinement, and researchers need to increase their activities through the Public Research and Regulation Initiative (PRRI) and other organizations to ensure that high-quality science informs CBD negotiations.

Steven H. Strauss is in the Department of Forest Ecosystems and Society, Oregon State University, Corvallis, Oregon, USA. Huimin Tan is at the Heinz School of Public Policy and Management at Carnegie Mellon University, Pittsburgh, Pennsylvania, USA. Wout Boerjan is at the Department of Plant Systems Biology, VIB and the Department of Plant Biotechnology and Genetics, Ghent University, Gent, Belgium. Roger Sedjo is at Resources for the Future, Washington DC, USA.
e-mail: steve.strauss@oregonstate.edu

A convention co-opted

Negotiated under the United Nations (UN) Environment Program, CBD was adopted in June 1992 and subsequently entered into force in December 1993. The CBD has been signed by 191 of the 192 members of the UN, making it one of the largest international treaties. The aim of the CBD is to promote the conservation and sustainable use of biodiversity, and the fair and equitable sharing of benefits from the use of genetic resources. Because transgenic organisms have the potential to affect biodiversity, special provisions of the CBD cover the use and trade in living modified organisms (LMOs, also known as genetically modified organisms; GMOs).

In 2000, the Cartagena Protocol on Biosafety was adopted based on the mandate in the CBD for a protocol on biosafety. It is supported by 147 members and its goal is to contribute to ensuring adequate protection, transfer and safe use in the field of GMOs that may have adverse effects on biodiversity. The focus of the Cartagena Protocol is transboundary movements, both intended and unintended. A main function of the Cartagena Protocol is to offer governments without national biosafety regulations a tool for informed decision making on the import of GMOs and to guide the development of national biosafety regulations.

Though the Cartagena Protocol has been ratified by almost all countries, many of the important details of the treaty are yet to be specified. They either are left up to the individual parties to implement as they see fit or have yet to be agreed upon. With respect to GMOs, this includes what is needed in risk assessments in specific cases; how to label GMOs during intercountry transfer; how to obtain public input; and how to deal with liability and redress for damages to biodiversity[1–4]. It also allows highly diverse interpretations of socioeconomic issues. Article 26 states that "the Parties…may take into account, consistent with their international obligations, socio-economic considerations arising from the impact of LMOs on the conservation and sustainable use of biological diversity, especially with regard to the value of biological diversity to indigenous and local communities." As discussed below, many of these uncertainties may present critical issues for forestry.

Will activists succeed in keeping the lid on transgenics tree research?

As biotechnologies are viewed in the CBD as having substantial potential benefits for biodiversity and sustainability, the goal of the Cartagena Protocol is not to prevent the use

Box 1 Diverse types and uses for transgenic trees

A main argument from scientists against broad bans or moratoria on all types of field studies with GM forest trees is that there is a large diversity of anticipated benefits and risks that need specific evaluation. The traits under study include wood chemistry, herbicide resistance, insect resistance, disease resistance, rate of growth, stature, salt tolerance, nutritional conditions, dormancy induction, onset of flowering, sterility, phytoremediation, cold tolerance, gene induction systems and rootability[33,34]. This diversity was underlined by the CBD's own background document prepared for the SBSTTA meeting in Rome in February 2008 (ref. 5,35), entitled "the potential environmental, cultural, and socio-economic impacts of genetically modified trees." In Annex 1 of that document, a long list of the kinds of potential environmental and socioeconomic and cultural impacts, both positive and negative, were enumerated and discussed. Similar lists of diverse benefits and impacts, as well as means for mitigation of undesired impacts, were provided in earlier reviews[31,36].

There is also a diversity of species being pursued in GM research. Frankenhuyzen and Beardmore identified 33 species of forest trees that had been successfully transformed and regenerated[31]. Although a majority of field trials have occurred in poplar (*Populus*) because of its status as a model organism for tree genomics and biotech, and most have occurred in the United States[34], field tests have also been conducted in a number of other tree species and geographies around the world[33,37]. Plantation trees predominate, with poplar leading (177 trials as of February 2008), followed by pine (129) and eucalypts (56)[38].

of transgenic or other biotechnologies but to guide their wise and safe use. But it is the risks, not the potential benefits, to biodiversity that have received the large majority of attention, mainly owing to the predominantly negative views of GMOs by some European Union (EU) member states and affiliated developing countries, and the prominence at the negotiations of nongovernmental organizations (NGOs), such as Greenpeace, that are conducting strong anti-GMO campaigns. The United States signed the CBD in June 1993 but has never ratified it[5], in part because of its hostile treatment of transgenic biotech—now a major feature of US agriculture and agricultural exports[6].

Only recently have GM trees and their role in forestry become an important feature of the anti-GMO campaigns. The rhetoric is often strong. For example, Anne Petermann of the Global Justice Ecology Project (http://www.globaljusticeecology.org/) stated that GM trees "…pose what many consider to be the most serious threat to the world's remaining native forests since the invention of the chainsaw"[7]. As with the broader GMO debate, the anti-GMO activists often cast the debate as people versus corporations. Petermann also wrote that there is "…mounting corporate pressure to deregulate GE [genetically engineered] trees so that they can be developed on a commercial scale for the future production of paper, biofuels, chemicals, plastics and other products"[7]. The benefits to broader society of these products, produced at reasonable costs on a potentially smaller land base than conventionally produced trees, are denied, disputed or ignored.

There is now a push for a moratorium or ban on all GM tree field tests through the Cartagena Protocol, including those which are strictly confined or done only for research. A similar effort to ban GM trees was mounted in conjunction with the negotiations surrounding the Clean Development Mechanism part of the Kyoto Protocol. Although the ban was not imposed, Clean Development Mechanism requirements for an environmental impact review and an executive board to provide oversight provides a means through which anti-GMO NGOs can continue to provide political influence. This is likely to make even research with GM trees very difficult in many countries[8].

As discussed below, the efforts against GM trees appear to be having a substantial influence on Cartagena Protocol recommendations and thus are likely to affect national and international regulations. Our aim is to examine the context for this campaign, and the extent to which it is consistent with scientific knowledge, the perspectives of scientific organizations and the views of the major environmental NGOs.

Biotechnologies and trees

A diverse array of biological technologies are being intensively pursued to support plantation forestry. These include clonal propagation, interspecific hybridization, use of exotic species, the use of a variety of molecular tools to intensify the selection of superior genotypes (DNA fingerprinting, genome mapping, gene identification and genome sequencing) and transformation[9]. However, of this diverse array of technologies, only transformation, defined by the use of direct modification and asexual insertion of DNA into organisms in the laboratory (that is, genetic engineering or modification), engenders attention from the CBD, strong government regulation and controversy over its use, even for research.

The goals for GM tree forestry are highly diverse, as are the locations, the species and the genes employed (**Box 1**). In addition to the use of genes from other species, genetic modification can involve changes of the expression of native genes to modify endogenous traits, such as wood structure, growth rate and tolerance of stress. Such activities have been increasing as knowledge of the genomes of trees increases, and genetic modification as a means to leverage genomic information is viewed as particularly important for trees versus annual crops because of the slow pace of tree breeding and the limited state of tree domestication[10]. Genomic information on major forestry species has increased dramatically in recent years. The entire genome sequence of the poplar tree (*Populus*, aspens and cottonwoods) was published by the US Department of Energy (DOE; Germantown, MD) in 2006 (ref. 11) and sequencing of the *Eucalyptus* genome, also by DOE, is currently underway. In addition to industrial purposes, efforts are underway to use recombinant technology to help rescue major tree species that have been devastated by exotic diseases, such as have occurred for chestnut and elm in the United States[12], to improve the efficiency of environmental cleanup[13] and to reduce the risks of ecological harm due to the spread of exotic tree varieties[14]. Products such as disease-resistant chestnut and elm should have direct benefits for promoting forest biodiversity by resurrecting key species that support many kinds of organisms in the ecosystems in which they occur.

Given the diversity of traits, species and environments under study, a case-by-case approach would seem to be the sensible way to proceed, and this basic approach is officially recognized in the Cartagena Protocol[2]. Annex III/6, under general principles governing risk assessment, states that "risk assessment should be carried out on a case-by-case basis. The required information may vary in nature and level of detail from case to case, depending on the LMO concerned, its intended use and the likely potential receiving environment." This principle fits well with the diversity of GM trees.

Views of scientific and environmental groups

Nonetheless, the activism against GM trees through the CBD has been against all forms of genetic modification, regardless of the goals or environmental benefits sought. This activism has also been in direct opposition to widespread scientific and professional opinion from around the world, including from ecologists (**Table 1**), that the trait, not the recombinant method,

should be the focus of ecological assessments. These views derive from some of the largest and oldest scientific and professional organizations knowledgeable on these issues, and are the result of intensive, high-level deliberations among diverse member scientists. A similar view was espoused in the Biosafety Regulation Sourcebook, created to help countries craft national regulations that are congruent with Cartagena Protocol rules and intentions: "The risk an organism or related activity may pose to the environment depends on the organism's properties and resulting interaction with the environment. This is the case regardless of whether those properties are the result of breeding technologies—either traditional techniques, or biotechnology—or 'natural' evolution. This fact has been and continues to be confirmed by leading international institutions including the OECD [Organization for Economic Co-operation and Development], FAO [Food and Agriculture Organization] and WHO [World Health Organization][15].

The majority of the major environmental NGOs also do not have policies that discriminate against all types of GMOs, with the notable exceptions of three large NGOs: Friends of the Earth, Greenpeace and the Sierra Club (**Table 2**). Although the three anti-GMO groups present themselves and their concerns as based on science, this disagreement on a fundamental principle that underlies scientific risk assessment suggests otherwise. In contrast, all of the major scientific organizations, and most of the major environmental NGOs, have not seen fit to promote indiscriminately anti-GM policies or campaigns.

Anti-GM tree campaigns

Active campaigns against GM trees through the CBD began in early 2004, with a coalition of small NGOs calling for a ban on GM trees

Table 1 Views of major scientific and professional societies on evaluation of genetically engineered crops and trees

Organization	Year created	Number of members	Total expenses at end of 2006	Quotation or position
American Medical Association[a]	1847	~278,000	$222,344,781[b]	"Federal regulatory oversight of agricultural biotechnology should continue to be science-based and guided by the characteristics of the plant, its intended use, and the environment into which it is to be introduced, not by the method used to produce it…" <http://www.ama-assn.org/ama/no-index/aboutama/13595.shtml>
American Council on Science and Health[c]	1977	NA	$1,845,871	"Current regulatory scrutiny, plus the excellent track record of GM food safety, gives us confidence that GM foods are rigorously scrutinized and that the technology is safe." <http://www.acsh.org/publications/pubID.289/pub_detail.asp>
American Society of Plant Biologists[c] (ASPB)	1924	~5,000	$5,418,347	"ASPB strongly endorses continued responsible development and science-based oversight of GE and all food production technologies and practices on a case-by-case basis." <http://www.aspb.org/publicaffairs/aspbgestatement.cfm>
American Seed Trade Association (ASTA)[c]	1883	~850 companies	$3,006,991	"ASTA strongly supports the safe use of new modern genetic methods in the continuing effort to improve crop varieties. The safety of crops modified by modern biotechnology is ensured through a most rigorous and comprehensive set of regulatory systems. The resulting varieties hold great promise for improving the food and feed supply of the world and promoting environmental sustainability, just as past accomplishments of plant breeders have benefited the world." <http://www.amseed.com/govt_statementsDetail.asp?id=43>
American Phytopathological Society[c] (APS)	1908	~5,000	$3,572,946	"(APS)… supports biotechnology as a means for improving plant health, food safety, and sustainable growth in plant productivity." <http://www.apsnet.org/media/ps/APS%20Biotech%20Statement.pdf>
Council for Agricultural Science and Technology[c]	1972	~38 scientific societies	$767,789	"Retain the current case-by-case safety assessment approach and continue to emphasize regulatory conditions carefully tailored to address risks identified for individual biotechnology-derived plant products. Agencies must maintain the flexibility to assure that rigorous, science-based safety assessments are conducted for each new product or product category." <http://www.castscience.org/displayNewsRelease.asp?idNewsRelease=118&display=1>
Ecological Society of America[c]	1978	~8,000	$3,609,200	"GEOs have the potential to play a positive role in sustainable agriculture, forestry, aquaculture, bioremediation, and environmental management, both in developed and developing countries." "We reaffirm that risk evaluations of GEOs should focus on the phenotype or product rather than the process…." <http://www.esa.org/pao/policyStatements/Statements/GeneticallyModifiedOrganisms.php>
Food and Agriculture Organization of the United Nations (FAO)[c]	1945	189 member nations	NA	"FAO supports a science-based evaluation system that would objectively determine the benefits and risks of each individual GMO. This calls for a cautious case-by-case approach to address legitimate concerns for the biosafety of each product or process prior to its release." <http://www.fao.org/biotech/stat.asp>
Genetics Society of America[c]	1985	~5,000	$3,123,807	"…it will be necessary to consider products on a "case-by-case" basis. In some cases, a GMO may not be different in any significant way from a classically bred organism." <http://www.genetics-gsa.org/pages/pp_benefits.shtml>
Institute of Food Technologists[a]	1939	NA	$15,934,326	"There is some evidence of overall improved environmental safety due to wider use of rDNA biotechnology. That is not to say that all rDNA biotechnology-derived products will be safe—they must be examined on a case-by-case basis before being commercialized." <http://members.ift.org/NR/rdonlyres/892A5152-5F08-4921-840C-03587DAA1F1B/0/iftreport_benefits.pdf>
International Society of African Scientists (ISAS)[c]	1982	NA	NA	"ISAS believes that agricultural biotechnology represents a major opportunity to enhance the production of food crops, cash crops, and other agricultural commodities in Africa, the Caribbean and other developing nations." <http://www.aspb.org/publicaffairs/agricultural/africanbiotech.cfm>

continued

Table 1 Views of major scientific and professional societies on evaluation of genetically engineered crops and trees (continued)

Organization	Year created	Number of members	Total expenses at end of 2006	Quotation or position
International Union of Forest Research Organizations [a]	1892	689 member organizations	NA	"The social discussion about risks vs. benefits of GMOs must move from a generic consideration of GMOs to the merits of modifying trees with specific traits to be used in specific environments and management regimes"[32].
National Agricultural Biotechnology Council (NABC)[a,c]	1988	>30 research-educational institutions in North America	NA	Whether or not a GEO requires bioconfinement "should be determined on a case-by-case basis...." <http://nabc.cals.cornell.edu/pubs/nabc_17/NABC17_complete.pdf> "...genetically improved products should be evaluated for safety on a case-by-case basis, utilizing all of the available information, including experience, to guide the assessment." <http://nabc.cals.cornell.edu/pubs/statement2000.pdf>
National Research Council[a]	1916	~6,000	Annual budget: ~$176 million	"...the product of genetic modification and selection should be the primary focus for making decisions about the environmental introduction of a plant... and not the process by which the products were obtained." <http://books.nap.edu/openbook.php?record_id=1431&page=67> "For purposes of decision support, the process of production should not enter into risk assessment." "The transgenic process present[s] no new categories of risk compared to conventional methods of crop improvement, but specific traits introduced by either of the approaches can pose unique risks." <http://books.nap.edu/openbook.php?record_id=10258&page=63> "Because both methods have the potential to produce organisms of high or low risk, the committee agrees that the properties of a genetically modified organism should be the focus of risk assessments, not the process by which it was produced." <http://books.nap.edu/openbook.php?record_id=9795&page=6>
Pontifical Academy of Sciences[c]	1603	~80 academicians	NA	"There is nothing intrinsic about genetic modification that would cause food products to be unsafe. Nevertheless, science and scientists are and should further be – employed to test the new strains of plants to determine whether they are safe for people and the environment, especially considering that current advances can now induce more rapid changes than was the case in the past." <http://www.vatican.va/roman_curia/pontifical_academies/acdscien/documents/sv%2099(5of5).pdf>
Society of American Foresters[c] (SAF)	1900	~18,000	$3,175,752	"SAF supports the continued evolution of federal regulations that affect forest tree biotechnology, particularly changes to make the regulations more focused on the products' safety and environmental impact, rather than on the process or method used to create them." <http://207.5.76.244/fp/documents/forest_tree_biotech.pdf>
Society of Toxicology[a]	1961	NA	$5,232,371	"...the potential adverse health effects arising from biotechnology-derived foods are not different in nature from those created by conventional breeding practices for plant, animal, or microbial enhancement." "...it is the food product itself, rather than the process through which it is made, that should be the focus of attention in assessing safety." <http://www.toxicology.org/ai/gm/GM_Food.asp>
The World Health Organization (WHO)[c]	1948	~191 member states	NA	"GM foods currently available on the international market have undergone risk assessments and are not likely to present risks for human health any more than their conventional counterparts. The potential risks associated with GMOs and GM foods should be assessed on a case-by-case basis, taking into account the characteristics of the GMO or the GM food and possible differences of the receiving environments." <http://www.who.int/foodsafety/publications/biotech/biotech_en.pdf>

[a]Obtained from reports and web pages that suggest a position on genetic engineering, not an official position statement. [b]Annual expenses at end of 2005. [c]Based on policy statement or position statement. NA, not available. Annual expenses for FAO, ISAS, NABC, IUFRO and WHO not available.

due to the high potential for wide dispersal of pollen and seed, which they argued goes against the basic tenets of the CBD[16]. This action appears to have been precipitated by the decision of the UN Framework Convention on Climate Change in December 2003 not to exclude GM trees in the Clean Development Mechanism. A small coalition against GM trees that formed late during those negotiations failed in getting them excluded from Clean Development Mechanism carbon accounting[8]. Later that year, GM trees were discussed at the fourth session of the UN Forum on Forests, where the anti-GMO NGOs present further argued for a global ban[17]. This action was continued during the second conference and meeting of the parties (COP-2, MOP-2) to the Cartagena Protocol on Biodiversity and the CBD during 2005 in Montreal, Canada[18], and again at COP-8 of the CBD in Curitiba, Brazil, where a request was made for the CBD to produce a report on the "potential environmental, cultural, and socio-economic impacts of genetically modified trees."

This report was first prepared in 2007 for a CBD-associated technical meeting in Montreal, then revised based on scientific reviews by PRRI (http://www.pubresreg.org/) scientists and others, and presented in final form at the CBD–Subsidiary Body on Scientific, Technical and Technological Advice (SBSTTA) meeting in Rome in 2008 (ref. 19). The document enumerates the many and diverse benefits and risks from the use of GM trees. Interestingly, nearly all of the same list of benefits and risks would apply to many forms of conventional breeding and associated intensive plantation forestry, but this is not explicitly discussed nor is a comparative risk assessment for GM trees specifically called for in the CBD. The risks touted against GM trees are discussed in **Box 2**.

The anti-GM tree campaign grew in numbers to include 137 organizations that were represented in Rome and Bonn in 2008 (refs. 20,21), most of them very small, but now including several that claim to represent indigenous peoples whose main concerns are land use, multinational corporations and the spread of intensive plantations generally, not GM trees specifically. Also included were the larger anti-GMO NGOs, including Greenpeace, Sierra Club, World Rainforest Movement and Friends of the Earth.

CBD recommendations

Two resolutions have been accepted by the Cartagena Protocol about GM trees, both urging precaution with respect to their study and use (**Box 3**). Both statements refer to the propensity for wide gene dispersal as a problem for the CBD with its attention to transboundary movement of LMOs. However, they do not discuss why this concern is singled out compared to dispersal of non-GM trees, which often are moved over long distances from their native ranges in breeding programs, can be the result of intensive selection for trait modification, and may include exotic species and hybrids that do not naturally exist in the regions where they are planted. They also do not address that many of the GM traits, such as those proposed for ease in processing biofuels[10], are expected to domesticate, rather than to invigorate trees, and thus should reduce risk of spread and associated impact on biodiversity compared to currently used trees. In other words, the resolutions do not provide any suggestions for comparative risk assessment to help make proportionate risk assessment decisions for the many different kinds and environmental values of GM trees, as is required in the Cartagena Protocol[2]. In fact, they suggest the opposite—that all GM trees as a class should be put through extreme scrutiny and be given special attention in risk assessments, in direct opposition to the case-by-case and product-not-process principles supported by scientific and environmental organizations (**Table 1**). The presumption of hazard from all types of GM trees is not based on a demonstration of generic hazard but is rather a presupposition of environmental risk that is without any scientific justification.

The recommendations also do not suggest that environmental and economic benefits be considered at all, nor do they consider that the process of stringent risk assessment, including the long duration of assessments suggested, is likely to foreclose substantial economic and environmental benefits. Finally, they do not point to the very large potential for GM approaches to reduce some of the risks of gene dispersal from conventional trees, for example, by engineering traits that reduce fertility of exotic or invasive species[22,23], and make no distinction between the very different risks of confined and small-scale field studies versus large-scale commercial releases.

It is clear that the groups strongly opposed to GM trees wish to regulate them out of existence directly, or achieve the same outcome by using the CBD's recommendations to direct national regulations toward requirements that are extremely costly or effectively impossible to meet—conditions that have already been developing in recent years. In the United States, the costs and requirements for permission to conduct multiple-year field trials has grown substantially in recent years owing to the requirement that all such tests obtain full permits from the US Department of Agriculture[24]. In the European Union, there have been only 18 trials of forest trees authorized over a period of 17 years, and attempts to do even short-term contained field studies of trees with modified versions of native genes can run into major legal and political snags. If there are no field demonstrations of value in model genotypes, there will be no further development of commercially useful GM varieties. A high, costly hurdle for field testing discourages investment both by industry and public sector organizations (**Box 4**).

Looking to the future

There is clearly a considerable potential for progress in tree improvement using GM technology given the advances in molecular biology and genomics of forest species. There is also a pressing need for innovations given the increased climatic stresses on plantation forests expected, and the importance of forests for biological materials, renewable energy, carbon sequestration, biological diversity and other

Table 2 Views of major US environmental NGOs on GM crops and trees

Organization	Founded	Expenses 2006	Position
Friends of the Earth	1970	$3,568,260	"In the case of the Convention on Biological Diversity, it is clear that GMOs in general and GM trees in particular, constitute a violation of the convention…" "We therefore call upon all governments, especially the Parties to the Framework Convention on Climate Change and its Kyoto Protocol, to ban the release of GM trees." <http://www.wrm.org.uy/subjects/GMTrees/text.pdf>
Greenpeace	1971	$15,556,440	"Greenpeace is opposed to the release of genetically engineered organisms into the environment at the present state of knowledge and calls for a ban on the release of transgenic trees. As an interim measure a global moratorium on commercial releases and on larger scale experimental releases is recommended." <http://www.genet-info.org/fileadmin/files/genet/GE_Trees/2006_GP_GETrees.pdf>
Int. Union Conservation Nature (IUCN)	1986	$902,112	"Research into GM applications should continue and indeed accelerate but with 'eyes wide open', assessing each GM application on a case-by-case basis." <http://cmsdata.iucn.org/downloads/ip_gmo_09_2007_1_.pdf>
Natural Resources Defense Council	1936	$63,774,845	"…we do not have an official position on [genetically engineered crops and trees]…" (J. Powers, NRDC New York Media Relations Director, personal communication on Nov. 19th, 2008).
Sierra Club	1960	$83,432,700[a]	"Sierra Club has taken no positions regarding genetic engineering done in labs or in indoor manufacturing of pharmaceuticals." "Sierra Club opposes the out-of-doors deployment of genetic technologies [GM trees]." <http://www.sierraclub.org/biotech/trees.asp>
The Nature Conservancy (TNC)	1951	$671,580,417	"…the Nature Conservancy does NOT have any specific policy or position on GMOs." (M. Tu, TNC, personal communication on July 24, 2008)
Union of Concerned Scientists	1969	$12,576,026	"Risks must be assessed case by case as new applications of genetic engineering are introduced." <http://www.ucsusa.org/food_and_environment/genetic_engineering/risks-of-genetic-engineering.html>
Worldwatch Institute	1974	$873,521	"The Worldwatch Institute has no position statement or policy with regard to genetically engineered crops and trees." (Robert Engelman, Worldwatch, personal communication, July 25, 2008). A recent paper published by a staff member suggests case-by-case consideration of merits and risks for specific products. <http://www.worldwatch.org/system/files/EP145B.pdf>

Obtained from reports and web pages that suggest position on genetic engineering, not an official position statement. [a]Annual expenses at end of 2005. NA, not available.

Box 2 What are the risks?

The objections of anti-GMO groups to transgenic trees generally fall into two categories (for a more detailed discussion of the concerns associated with transgenic trees, see ref. 39): risks implicit to the use of recombinant technology and risks associated with specific GM traits under development.

With respect to generic concerns related to recombinant technology, the mutagenesis that accompanies the process is often portrayed by anti-GMO groups as unacceptably large. But molecular variation induced by genetic modification pales when compared with the level of genetic diversity among conventional varieties[40–43]. In a study of maize diversity in the absence of genetic modification, Morgante *et al.*[44] conclude that "the maize genome is in constant flux, as transposable elements continue to change both the genic and nongenic fractions of the genome, profoundly affecting genetic diversity." For trees, the variable effects of different gene insertions are often cited, yet the unpredictability associated with common methods of tree breeding, such as interspecific hybridization and long-distance geographic transfers, are ignored.

As to risks related to traits, such as lower lignin composition or fertility reduction, the scientific consensus is that such traits are not threats to wild forests, as often claimed, because they tend to reduce fitness, impeding their own spread. In addition, trees modified with these genes would have to pass many years of field tests for health, stability and adaptability before large-scale use in plantations, making large-scale plantation failure unlikely. What's more, the changes in ecological chemistry imparted by GM traits such as these tend to be modest compared with normal silvicultural manipulations and intensive breeding (e.g., planting density, vegetation control, shifts in planted tree species and interspecific hybridization), and there are many ways to mitigate impacts by stand-level and habitat management, such as the use of buffer strips, mosaic plantings or rotations with diverse species or genotypes. Such traits as herbicide tolerance will be accepted or rejected on the basis of how their use affects vegetation control and biological diversity both inside and outside of managed forests. Finally, horizontal gene flow, including that of selectable marker genes, has never been shown to occur in nature at a rate that is of ecological concern, nor are there reasons to expect that such transfers could create significant novelties in comparison to the extraordinary diversity of microbial genomes and antibiotic resistance genes[45,46].

Perhaps the most credible science-based concerns about GM trees relate to their potential for wide dispersal of seeds and pollen when they are allowed to flower. Although several forest tree species, including poplar, can also spread vegetatively, this way of propagation tends to be much more localized, much less frequent and can be far better controlled when required in regulations or in commercial practice. The strong concerns about gene dispersal are illustrated by these comments from Petermann[15] in her description of issues at the recent CBD meetings in Bonn, Germany: "The incidents of contamination [with GE [genetically engineered] agricultural crops] show that gene escape and GE contamination cannot be prevented once GE crops are released. This in turn suggests that the widespread planting of GE trees would over time lead to a persistent contamination of the world's native forests, with disruptive ecological consequences."

There is wide agreement from scientists that until very strong containment genes are developed, socially accepted and their efficiency verified in the field, some level of gene dispersal—either from pollen, seeds or vegetative propagules—is certain in most forestry species[14,39,47–49]. Moreover, the distances over which dispersal can occur are large, on the order of kilometers or more. This is mostly a consequence of the potential for long distance movement of pollen by wind and pollinating insects, and to a lesser but still considerable extent owing to movement by seeds. The latter can occur when seeds are very small, subject to movement by major storm systems, or are dispersed by animals such as birds. The limited level of domestication of most tree species contributes to this concern, as propagules are generally fit enough to survive in wild or feral environments.

However, the biological significance of this gene dispersal needs to be put into perspective. First, adventitious presence at a low level is also often prevalent with non-GM crops and trees and usually does not create significant ecological problems (it is an ongoing fact of agriculture and forestry using selectively bred and/or exotic genotypes). Second, compared with the diversity of wild forests, very few GM species are under commercial development that are sexually compatible with wild forests, or will be used in or very near to wild forests, and thus it will be extremely rare that transgenes could introgress into wild tree genomes to a significant degree, and thus become common in wild ecosystems. The area planted with GM forest trees is likely to remain relatively small; forest plantations comprise only ~5% of the world's forest cover[50]. Third, there may be potential benefits for wild tree species from some kinds of GM trees; for example, a wild tree might benefit by acquiring a trait enhancing stress resistance and thus acquire resilience in the face of new forms of biotic or abiotic stresses, perhaps brought on by rapid climate change[12,14]. Fourth, the quantitative amount of admixture may be so low as to be trivial in ecological impact, owing to distance and dilution from extensive wild forests, as a result of intentional use of (even imperfect) containment genes, and from the selective disadvantage imparted when domesticating traits are conferred[51]. Fifth, although concerns have been raised about the effects of containment genes on biodiversity were floral/fruit organs to be altered or removed, by appropriate technology selection (e.g., to selectively target tissues and gametes), and by the rational deployment on the landscape that is already common in plantation forestry, the impacts on biological diversity can be responsibly managed (references and discussion in ref. 14). And finally, it is not clear that GM-imparted traits have the capability to substantially and sustainably improve fitness such that there is sufficient spread and persistence to produce "disruptive ecological consequences"[6], especially given the continued high levels of environmental change and rapid pest evolution. In sum, as a result of all of these factors, most scientists emphasize not whether some gene dispersal will occur; they assume some level may occur for the foreseeable future, but focus on what the extent might be (how frequent, over what distance), and if any substantial adverse consequences (ecological, economic) are likely compared with the expected level of environmental change from other sources, and how these alterations compare to the benefits brought by the GM varieties.

Thus far, however, very few field studies have been conducted that are on the scale needed for useful ecological inferences, in large part because of the regulatory restrictions in doing so[31]. Cartagena Protocol recommendations appear to be putting national regulatory policies on a path toward making such research even more difficult, and for many purposes, impossible to carry out.

ecological services. Even so, progress in translating genomic science into application requires field studies and ultimately decisions from societies about what kinds of innovations are reasonable in the environment at the research and application stages. Unfortunately, applying the 'precautionary approach' or the much vaguer and politically malleable 'precautionary principle'[25] to GM trees, though recommended in recent CBD meetings and well-intentioned in its original goals, appears to confound progress with transgenic tree research.

The precautionary principle has been interpreted in a myriad of ways, depending on the political interests of the parties involved[26,27]. An excess of precaution can lead to calls for extensive and long-term studies of trivial biological issues compared with conventional breeding and silviculture, with costs so great as to effectively halt further investment by the private and public sector. As discussed by Kinderlerer, "A problem with the debate on precaution is that the absence of consensus within the scientific community, especially where weight is attributed equally to all scientists, provides ammunition for those who for many reasons wish to argue against the development of modern biotechnology"[4]. The very promise of novelty and innovation provided by modern biotech, with its new types of genetic innovations, becomes reasons to avoid all development under one view of strict precaution. Under an equally legitimate view, however, precaution demands that we pursue a wide array of options about future forestry and natural resource supplies, and because of their wide potential benefits, it would seem to provide a compelling reason to move forward with transgenic forest biotechnologies.

How the major uncertainties about the details of required risk assessments, unintended transboundary movement, and liability and redress will be worked out present special concerns for research on GM trees. The potential long distances of gene dispersal with trees make strict containment within national boundaries difficult in many places—especially when considered over many planting cycles. The responsibilities of users of LMOs under Article 17 of the Cartagena Protocol that governs unintended transboundary movements are unclear. The long life cycles of GM trees make empirical risk assessment studies of ecological effects slow and costly. It also remains unclear whether growers of GM trees will be held liable under Article 27 of the Cartagena Protocol for unintended dispersal, and how the socioeconomic impact provisions would encompass losses of income from such spread. The Cartagena Protocol was created to address impacts on biodiversity from new traits that result from use of LMOs, not the simple presence of GM DNA[15]. Yet, organizations such as the Forest Stewardship Council, a major international certifier of 'green' and socially responsible forestry and forest products, treats all GM trees, even contained and short-term field research with obvious environmental goals, as a major violation that would void certification. Its treatment of contamination by pollen, seeds or vegetative propagules of a non-GM certified forest or product, and the CBD consideration of such actions, are unclear[4,28]. It is also unclear what parties would be liable, and whether this would include growers, seed companies or regulators in government bodies that authorize field uses. The latter risk is of particular concern given the proliferation, yet lack of technical capacity, to adequately administer biosafety regulatory agencies in many countries. A report from the UN University Institute of Advanced Studies[29] concludes that: "there remains a significant lack of capacity in many developing countries…[and a] country that lacks capacity is more likely to bring in very restrictive systems in order to counterbalance its deficiencies….[Thus, the] lack of an effective biosafety regime undermines the potential for developing countries to consider the role of biotechnology in critical areas such as addressing climate change…"

Until recently, public sector inputs about biotech at the CBD have largely come from anti-GMO–oriented NGOs[30]. Only in the past few years have public sector scientists had a large presence, mainly through the PRRI (**Box 5**). The PRRI organizes and brings scientists to the negotiations to explain the value of transgenic biotech for public sector research and for broad public benefit, and to correct the biased, incomplete or false statements about LMOs that are frequently and loudly made by anti-GMO

Box 3 Precautionary approach and principle

Recent resolutions on GM trees taken at CBD/Cartagena Protocol associated meetings, if interpreted literally and used to guide national biosafety regulations, would clearly have a major chilling effect on field research on opportunities for commercial development of GM trees. The decision on GM Trees made at COP-8 in Brazil states in part: "The Conference of the Parties, recognizing the uncertainties related to the potential environmental and socio-economic impacts, including long-term and trans-boundary impacts, of genetically modified trees on global forest biological diversity, as well as on the livelihoods of indigenous and local communities, and given the absence of reliable data and of capacity in some countries to undertake risk assessments and to evaluate those potential impacts…recommends parties to take a precautionary approach when addressing the issue of genetically modified trees "[52].

The relevant section from the recent COP meeting in Bonn in 2008 (ref. 53) states that the CBD urges parties to "[1] reaffirm the need to take a precautionary approach when addressing the issue of genetically modified trees. Principle 15 of the Rio Declaration on Environment and Development [states that] In order to protect the environment, the precautionary approach shall be widely applied by States according to their capabilities. Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation. [2] Authorize the release of genetically modified trees only after completion of studies in containment, including in greenhouse and confined field trials, in accordance with the national legislation where existent, addressing long-term effects as well as thorough, comprehensive, science-based and transparent risk assessments to avoid possible negative environmental impacts on forest biological diversity (where applicable, risks such as cross-pollination and spreading of seeds should be specifically addressed). [3] Consider the potential socio-economic impacts of genetically modified trees as well as their potential impact on the livelihoods of indigenous and local communities. [4] Acknowledge the entitlement of Parties, in accordance with their domestic legislation, to suspend the release of genetically modified trees, in particular where risk assessment so advises or where adequate capacities to undertake such assessment is not available. [5] Further engage to develop risk assessment criteria specifically for genetically modified trees…."

These recommendations impose obstacles that may be insurmountable for field research in forest biotech. Although the precautionary approach appears to be a less vague guideline than the precautionary principle, its meaning and implementation are still open to wide variation in interpretation. The recommendation from the Bonn meeting to address "long-term effects," even though GM trees are not generally allowed to flower or reproduce in the field under "containment," except in exceptional and often unaffordable conditions of isolation, appears to impose a Catch-22, meaning that there is no way for most countries and organizations to move forward. Given the enormous diversity in GM traits, benefits and biological safety, there is no scientific rationale that can support such indiscriminate and draconian restrictions.

Box 4 The importance of field trials

Whereas anti-GM tree activists see no field research as safe, field studies can be conducted with a very high degree of biological safety and genetic containment and are essential for research to proceed beyond the basics. In most tree species, it is considered a simple matter to conduct highly contained field studies of several years' duration because during that time frame most forest trees have not yet begun to flower, or the flowers are few, close to the ground, and most or all can be readily removed or bagged. Thus, the risk of spread by pollen and seed is low—arguably lower than for many annual crops for which flowering and seed/fruit production happens rapidly, and for which seed/fruit production (rather than wood production) is essential to the goals of the trial. In addition, regulatory authorities generally require monitoring for pollen, seed, seedling and vegetative spread from field trials, and where spread from these processes is a risk, they require removal of flowers before maturity and gamete release. They also require monitoring for, and destruction of, seedling and vegetative propagules up to several years after the trial is complete. In support of field testing, the Global Industry Coalition concerned with regulation of transgenic trees stated that 700 field trials of transgenic trees had been conducted worldwide, without any harmful effects on biological diversity identified[54].

Without field studies, the economic value of newly imparted traits in comparison to conventionally bred trees, and the extent of ecological impact, cannot be adequately assessed[31,38,55]. Indeed, the need for carefully conducted field experiments has been emphasized for other types of crops both to develop useful models of ecophysiology[56] and to enable transgenic or molecular marker–based improvement of complex traits, such as drought tolerance[57]. Even with physiological perturbations as striking as those from elevated CO_2—and for which there have already been abundant field studies—a recent article emphasized the critical need for more field trials to enable realistic assessments of the ecological effects of rising CO_2 levels. Soil and herbivore communities are vastly more complex than can be effectively simulated in a microcosm over a short time period, and plants in the field experience highly variable and strong fluctuations in climate and biotic pressures that materially change patterns of gene expression compared to the simple stresses imposed in controlled, greenhouse experiments[58]. There are many anecdotal stories of places where field and lab or greenhouse results strikingly disagree, but few of these are published. Two that we are aware of for transgenic trees include the *4CL* antisense gene and a *LEAFY* promoter::barnase sterility gene, both in poplar trees. Poplars transgenic for *4CL* exhibited double the rate of growth of controls in one small greenhouse study[59] but in a randomized field study by our group showed only a negligible growth advantage or poorer growth (S.S. *et al.*, unpublished data). In the other case, the floral sterility transgene had no effect on tree growth rate and health in a careful greenhouse trial but was later found to be strongly deleterious in the field[60]. As discussed in the text, there have been many hundreds of field trials already conducted without report of an adverse environmental impact—suggesting that field data can be gathered without significant environmental risk.

NGOs and parties. To the extent that PRRI continues to find scientists that are willing to spend time at these political fora, the CBD will be able to hear a more balanced view of the scientific issues. Similar concerns that nonscientific agendas have become prominent at the CBD, even at the purportedly technical SBSTTA meetings, and that few actual scientists are therefore willing to attend them, also pertains to many other issues under discussion at CBD meetings[30]. Even so, it is difficult to find scientists that are willing to take part as the quality of scientific discussions tend to be extremely low and highly combative, and so are often demoralizing to them. It also takes a considerable effort by PRRI to fund the high costs of international travel for the scientists. Another problem is that the pool of public sector scientists working on transgenic approaches to breeding, and who are thus interested in advocating for sound regulations, appears to continue to decline as a result of the huge regulatory costs and market obstacles to commercial use of the derived varieties. As pointed out by Kinderlerer, "Although it is likely that most of the almost 200 countries that are members of the Convention on Biological Diversity are using modern biotechnology in their research institutions and universities, few are considering the commercialization of products that are likely to be the subject of transboundary movement as defined in the Cartagena Protocol on Biosafety"[4]. Until more public sector scientists believe that GM trees can be used in field research without undue regulatory burden or risk from vandalism, and that they can pass regulatory approval and lead to

Box 5 The Public Research and Regulation Initiative

National regulations are strongly influenced by international agreements, such as the Cartagena Protocol on Biosafety. However, during the development of international agreements the public research sector, which counts tens of thousands researchers in several thousand research institutes in developing and developed countries, had not been represented in an organized way[30]. In 2004, the Public Research and Regulation Initiative was established with the objective of providing public researchers involved in modern biotech a forum through which they are informed about, and can be involved in, relevant international discussions such as the Meetings of the Parties to the Cartagena Protocol. The goal of participation in such meetings is to inform the negotiators about the objectives and progress of public research in modern biotech, to bring high quality science to the negotiations.

The PRRI has taken a stand on GM trees, and issued the following statement at CBD meetings: "Classical breeding has made major contributions to improving the productivity of plantation forests. However, the current challenges caused by population growth, climate change and fossil energy shortage cannot be met by conventional breeding alone. To meet our trans-generation responsibility, we have to find solutions today. We strongly believe that modern biotechnology, including genetic modification, can contribute significantly to finding solutions in these areas. Given the large potential environmental and socio-economic benefits of GM trees and the extensive safety record of the hundreds of field trials with GM trees conducted worldwide, there is no scientific justification for a blanket suspension of releases of GM trees. Field research is, in fact, the only way to get realistic answers to the many questions that were so well developed in the background document on GM trees."

useful products for society, there is unlikely to be the critical mass of scientists willing to take part in CBD and other regulatory negotiations.

Of most immediate concern are the increasingly strict regulations that impede or preclude even field research, and thus the increased foreclosure of opportunities for commercial development. These restrictions on research also provide a signal to companies and public sector institutions that investments in GM tree research are not likely to ultimately be usable or profitable. With respect to scientific concerns, these restrictions also make it nearly impossible to answer the questions that regulators want answered about comparative safety. As discussed by Frankenhuyzen and Beardmore[31] in their extensive review of GM trees, the: "…evaluation of … risks is confounded by the long life span of trees, and by limitations of extrapolating results from small-scale studies to larger-scale plantations. Issues that are central to safe deployment can only be addressed by permitting medium- to large-scale release of transgenic trees over a full rotation. Current regulations restricting field releases of all transgenes in both time and space need to be replaced with regulations that recognize different levels of risk (as determined by the origin of the transgene, its impact on reproductive fitness, and nontarget impacts), and consider potential benefits, and assign a commensurate level of confinement." Ecologists and biotechnologists largely agree that without field studies, science-based regulatory decisions are not possible. By recommending increased stringency (precaution) for all kinds of GM trees, the CBD is making the very studies needed to resolve regulatory quandries increasingly difficult and in many places impossible. The effective prohibition on all types of GM trees that negotiations surrounding the CBD recommendations are helping to promote is clearly against both its spirit and intent.

ACKNOWLEDGMENTS

The authors thank Julian Kinderlerer, Val Giddings, Piet van der Meer, Rachel Lattimore and Leslie Pearson for their helpful comments on the manuscript.

COMPETING INTERESTS STATEMENT

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturebiotechnology/.

1. Kalaitzandonakes, N. *Regulation* **29**, 18–25 (2006).
2. Jaffe, G. *J. Public Aff.* **5**, 299–311 (2005).
3. Watanabe, K.N. *et al. Nat. Biotechnol.* **22**, 1207–1208 (2004).
4. Kinderlerer, J. *Collection Biosafety Rev.* **4**, 12–65 (2008).
5. Convention on Biological Diversity. *Parties to Convention on Biodiversity.* <http://www.cbd.int/convention/parties/list/> (2009).
6. International Service for the Acquisition of Agri-biotech Applications. *Global Status of Commercialized Biotech/GM Crops: 2007* <http://www.isaaa.org/Resources/Publications/briefs/37/executivesummary/default.html>
7. Petermann, A. *The International Status of Genetically Modified Trees (7/25/05).* <http://www.worldagroforestrycentre.org/downloads/International%20Status%20of%20GE%20Trees.pdf>
8. Schwartz, J. *N. Y. Univ. Environ. Law J.* **14**, 421–480 (2006).
9. Grattapaglia, D. & Kirst, M. *New Phytol.* **179**, 911–929 (2008).
10. Boerjan, W. *Curr. Opin. Biotechnol.* **16**, 159–166 (2005).
11. Tuskan, G.A. *et al. Science* **313**, 1596–1604 (2006).
12. Merkle, S.A. *et al. Tree Genet. Genomes* **3**, 111–118 (2007).
13. Doty, S.L. *et al. Proc. Natl. Acad. Sci. USA* **104**, 16816–16821 (2007).
14. Brunner, A.M. *et al. Tree Genet. Genomes* **3**, 75–100 (2007).
15. Arent Fox & International Environmental Resources. *Biosafety Regulation Sourcebook.* <http://www.arentfox.com/modelbiosafetyact.pdf> (2006).
16. Institute of Science in Society. *GM Trees Lost in China's Forests.* <http://www.i-sis.org.uk/GMTGL.php> (2005).
17. Humphreys, D. *et al. Biotechnology in the forest? Policy options on research on GM trees.* <http://www.genet-info.org/fileadmin/files/genet/GE_Trees/2005_EFI_GETrees.pdf> (2005).
18. Lang, C. The Convention on Biodiversity, GM trees and paper consumption. *World Rainforest Movement Bulletin*, January 2008. <http://www.wrm.org.uy/bulletin/126/viewpoint.html#Convention>
19. Subsidiary Body On Scientific, Technical And Technological Advice. The Potential Environmental, Cultural and Socio-economic Impacts of Genetically Modified Trees. <http://www.cbd.int/doc/meetings/sbstta/sbstta-13/information/sbstta-13-inf-06-en.pdf> (2008).
20. Global Justice Ecology Project. *GJEP/ STOP GE Trees Campaign Interventions at UN Biodiversity Convention.* <http://www.globaljusticeecology.org/stopgetrees_news.php?ID=132> (2008).
21. Convention on Biological Diversity Alliance. *No GE Trees, No Case By Case.* http://www.cbdalliance.org/sbstta-13/21_4.pdf (2008).
22. Hoenicka, H. & Fladung, M. *Trees Struct. Funct.* **20**, 131–144 (2006).
23. Richardson, D.M. & Petit, R.J. in *Landscapes, Genomics and Transgenic Conifers* (ed. Williams, C.G.) 169–188 (Springer Netherlands, 2006).
24. United States Department of Agriculture. *Policy statement regarding releases of perennials under notification.* <http://www.aphis.usda.gov/brs/pdf/BRS_Perennials_Statement.pdf> (2008).
25. Morris, J. *Toxicology* **181–182**, 127–130 (2002).
26. Marchant, G.E. & Mossman, K.L. *Arbitrary and Capricious: the Precautionary Principle in the European Union Courts.* <http://www.policynetwork.net/uploaded/pdf/Arbitrary-web.pdf> (2005).
27. Conko, G. *Transgenic Res.* **12**, 639–647 (2003).
28. Strauss, S. *et al. J. For.* **99**, 4–7 (2001).
29. Johnston, S. *et al. Internationally Funded Training in Biosafety And Biotechnology - Is It Bridging the Biotech Divide?* <http://www.ias.unu.edu/sub_page.aspx?catID=111&ddlID=673> (2008)
30. De Greef, W. *Nat. Biotechnol.* **22**, 811–812 (2004).
31. Frankenhuyzen, K.v. & Beardmore, T. *Can. J. For. Res.* **34**, 1163–1180 (2004).
32. Strauss, S. *et al. Nat. Biotechnol.* **17**, 1145 (1999).
33. Food and Agriculture Organization of the United Nations. *Preliminary Review of Biotechnology in Forestry, Including Genetic Modification* (Forest Resources Division FAO, Rome; 2004).
34. Information Systems for Biotechnology. *Field Trial Database.* <http://www.isb.vt.edu/cfdocs/fieldtests1.cfm> (2008).
35. United Nations Environment Program. *The Potential Environmental, Cultural and Socioeconomic Impacts of Genetically Modified Trees.* UNEP/CBD/SBSTTA/13/1 <http://www.cbd.int/doc/meetings/sbstta/sbstta-13/information/sbstta-13-inf-06-en.pdf> (2008).
36. Strauss, S.H. *et al. J. For.* **99**, 4–7 (2001a).
37. Robischon, M. Field Trials with Transgenic Trees - State of the Art and Developments. (Springer, Berlin Heindelberg, 2006).
38. Safety, G.M.O. *Deliberate Release of Genetically Modified Trees.* <http://www.gmo-safety.eu/en/wood/poplar/54.docu.html> (2007).
39. Farnum, P. *et al. Tree Genet. Genomes* **3**, 119–133 (2007).
40. Cheng, K.C. *et al. J. Agric. Food Chem.* **56**, 3057–3067 (2008).
41. Batista, R. *et al. Proc. Nat. Acad. Sci. USA* **105**, 3640–3645 (2008).
42. Baudo, M.M. *et al. Plant Biotechnol.* **4**, 369–380 (2006).
43. Catchpole, G.S. *et al. Proc. Natl. Acad. Sci. USA* **102**, 14458–14462 (2005).
44. Morgante, M. *et al. Nat. Genet.* **37**, 997–1002 (2005).
45. Ramessar, K. *et al. Transgenic Res.* **16**, 261–280 (2007).
46. Martínez, J.L. *Science* **321**, 365–367 (2008).
47. Slavov, G.T. *et al.* Gene Flow in Forest Trees: Gene Migration Patterns and Landscape Modelling of Transgene Dispersal in Hybrid Poplar. (CABI Publishing, Cambridge, MA, USA; 2004).
48. Smouse, P.E. *et al. Tree Genet. Genomes* **3**, 141–152 (2007).
49. Linacre, N.A. & Adesb, P.K. *Ecol. Modell.* **179**, 247–257 (2004).
50. El-Lakany, M.H. *Unasylva* **217**, 45–47 (2004).
51. Strauss, S.H. *Science* **300**, 61–62 (2003).
52. COP-CBD. *Forest Biological Diversity: Implementation of the Programme of Work.* COP8 Decision VIII/19 <https://www.cbd.int/decision/cop/?id=11033> (2008).
53. COP-CBD. *Report of the Subsidiary Body on Scientific, Technical and Technological Advice on the Work of its Thirteenth Meeting.* UNEP/CBD/COP/9/L.33: CBD 2008 <http://www.cbd.int/doc/meetings/cop/cop-09/official/cop-09-03-en.pdf> (2008).
54. Global Industry Coalition. *Compilation of Environmental Risk Assessment Guidance: Transgenic Trees.* <http://www.croplife.org/library/documents/Biotech/Biosafety%20Protocol/risk%20assessment/FINAL%20compilation%20of%20ERA%20for%20transgenic%20trees.doc> (2007)
55. Brunner, A.M., Busov, V.B. & Strauss, S.H. *Trends Plant Sci.* **9**, 49–56 (2004).
56. Yin, X. & Struik, P. *New Phytol.* **179**, 629–642 (2008).
57. Campos, H. *et al. Field Crops Res.* **90**, 19–34 (2004).
58. Laurentius, A.C. *et al. Science* **320**, 880–881 (2008).
59. Hu, W. *et al. Nat. Biotechnol.* **17**, 808–812 (1999).
60. Wei, H. *et al. Mol. Breed.* **19**, 69–85 (2007).

Masters of their universe

Randy Osborne & Laura DeFrancesco

Nature Biotechnology **talks to some of the leading characters behind the Genentech legend.**

Genentech—the biotech venture that launched a thousand companies—is no longer its own master. In March, majority stakeholder Roche reached an agreement with the South San Francisco, California–based company under which the Swiss drug maker would take over the biotech for ~$46.8 billion. "Nothing will change" was the mantra from Genentech, but everybody knew it would. A management shake-up the following month unseated 14-year CEO Arthur Levinson, credited with steering Genentech's uniquely entrepreneurial culture. He was named chair of Genentech's new board of directors, charged with overseeing the firm's integration with Roche. Another jolt came in May, when Susan Desmond-Hellman, Genentech's president of product development, left to become chancellor at the University of California, San Francisco (UCSF).

But many remember those first years when a small team of bright, intellectually disciplined young scientists—often rowdy and personally eccentric people—got the company up and running. We caught up with a few of those pioneers to talk about that era, their time and how they felt leading the charge.

Not your ordinary set of beach bums. Genentech's brain trust circa 1982 (from left to right): Dennis Kleid (still at Genentech, the longest-term employee); David Goeddel (founder and CEO of Tularik, (now Amgen), currently in venture capital); Art Levinson (until April, CEO of Genentech, now chairman of the board); Herb Heyneker (serial entrepreneur now in venture capital); Peter Seeburg (professor at the Max Planck in Heidelberg); Dick Lawn (CSO at CV Therapeutics, now part of Gilead); and Axel Ulrich (Director of Molecular Biology at Max Planck Martinsried, founded Sugen, Axxima, U3Pharma).

Art Levinson

CEO Arthur Levinson's legendary time at Genentech began, as many legends do, inauspiciously in 1980. "I was finishing up my post-doctorate and getting antsy," says Levinson, fascinated by biology from an early age. "My wife was finishing her master's degree at the University of California Berkeley. I didn't want to do a fourth year of post-doc, and [Herbert Boyer, cofounder of Genentech] had just hired five new hotshots. I knew some of these folks. I thought I could spend a year and a half at Genentech, and learn how to clone genes."

He was still there three decades later. "It was something I never thought about," Levinson explains. "If I'm happy in the moment, then I don't worry about the next moment. I'm not saying that's good or bad, that's just the way I'm wired." Given "four or five bad days in a row," he might have considered moving on. Those days never came.

David Goeddel, one of the hotshots already recruited by Boyer, says Levinson "didn't take very long" to catch on. "He became very focused. I gave him his first reagents and advised him how to do some cloning. He figured it out quicker than anyone, and was on his way. Within a year, he was making direct contributions to the company."

Levinson recognizes, as does practically everyone in the industry, that Genentech's unprecedented success grew out of great effort by brilliant young scientists, helped by a nascent industry that still tolerated long-shot experimentation that often went nowhere. But

Randy Osborne is a freelance writer in Mill Valley, California and Laura DeFrancesco is Senior Editor at Nature Biotechnology.

he didn't expect it. "It's so easy to reconstruct history and pretend that these folks thought they were going to revolutionize everything," he says. The vast accomplishments seemed "completely random. There was no expectation of that at all."

Genentech's rowdy core of scientists vaguely sensed that "given a bit of luck, we might be able to do something terrific," Levinson says. What they all understood more acutely was that, if something terrific did happen, it would more likely take place within a place like Genentech, where "you didn't really know that you weren't in academia," he says. With the exception of Merck, "the big drug companies were moribund—it was like going to a morgue or something. People would show up at 9 and turn off the lights at 4:30. We needed to compete not with the miserable drug companies, but with MIT and Stanford." For the most part, too, the pharma firms hired "third and fourth-rate scientists," Levinson says—unlike Merck, which recruited "really great, really smart people who worked their butts off."

Though unspoken, there was a special urgency in the early days, a realization that "the place might fall apart in two years" if results failed to appear, he recollects. "We could all come to work and publish great papers and do great science, but we knew if we couldn't turn it into medicines that could help people, the whole thing was going to end." And because the Genentech way had never been tried, nobody could be certain about the outcome of such an approach. "It took 25 years to get the answer," Levinson says, adding that he was "only satisfied five or ten years ago that it was a good business model." He's satisfied now.

Herb Heyneker

Herb Heyneker's route to Genentech was through Boyer's lab, where he spent several years as a post-doc, after which he returned to his native Holland to practice the craft he had learned in the States. But the climate was not right for using the fancy tools, as it was post-Asilomar, and the Dutch were not ready to embrace the new technology. Boyer and Robert Swanson, cofounder with Boyer of Genentech, flew to Holland to invite Herb back, and back he came, enzymes in hand, literally. Herb's first of many contributions to the early work was the know-how to stitch DNA together, and on the first trip back to the States, he carried on the plane with him a thermos of DNA ligase.

You can't help but think that those early days at Genentech were special, as the word Herb most often uses in talking about it is "fantastic," notwithstanding the risk they all were taking. "We were babes in the woods. We were too naive to think it wouldn't work," he says.

Being the first to do something meant they left themselves open to criticism, and some, coming from their former academic colleagues, could be quite harsh. They were accused of "selling out to industry," and it was widely expected that secrecy would prevail. "In the early days, people looked at us with suspicion. At scientific conferences, we felt funny. But then we published in the same journals, the quality of research was good, we earned the respect of our peers. Things got better," says Heyneker. He, in particular, got the reputation of being open and forthcoming. According to Boyer, "Herb talked to everyone. He would tell the guy on the corner waiting for a bus what he was doing"[1].

Heyneker left Genentech after seven years to take on new challenges at a joint venture between Genentech and Corning Glass, Genencor. He felt at the time that he was still a part of Genentech, but the company didn't see it that way and rather soon thereafter, dumped Genencor's shares. After some successes in producing industrial enzymes (we have Herb to thank for high fructose corn syrup), he moved on, spent a few years at Stanford but then returned to entrepreneurial life, founding several companies (GlycoGen, GenPharm, ProtoGen and finally Eos) before joining the venture capital firm Abingworth. He now consults and sits on a number of boards.

Dave Goeddel

Credited by Boyer as responsible for more biologic drugs than anyone else[1], San Diego-born David Goeddel arrived with a bachelor's degree in chemistry and "exceptional" energy and drive, says City of Hope researcher Arthur Riggs, member of the early team that cloned somatostatin and insulin. His prankster sense of humor was hard to miss, too. Some probably wished they had.

"We were extremely competitive with the outside world," Goeddel remembers, but there was internal competition as well. "We had some issues with authorship on papers. Some guys got pretty hot with each other, and wouldn't speak for a few months. The nature of the projects demanded that [spirit]. Almost everything we did early on was publishable work in the top journals. It had to be up to that quality."

The drive to win extended everywhere. Staffers held races to determine who could load a gel or make a DNA prep faster. Then they would pause for a game of Nerf basketball or a standing broad jump contest—in the lab.

Practical jokes provided even more relief, and Goeddel became known as a grand master.

He led the caper in which 'radioactive' water from a research bin was spilled on the desk of the company's general counsel. After he wiped up the spill, he was told of the alleged danger, which was verified by passing a rigged Geiger counter over his body. The needle went berserk. "We had him take his suit off, and he was in the lab scrubbing down his face and chest in the lab," Goeddel says.

Another staffer was alerted by a call supposedly from his dentist's office, saying that serious carcinogens had been found in his teeth during his most recent exam. But, he explained, he had an important meeting with management that day, where he was scheduled to outline a pet project. "No, you have to come in right away," the caller insisted. "We have a team of six people who are going to work on you." He went, confronting a befuddled dentist and missing the presentation he had waited months for.

"There was one of that magnitude almost weekly," Goeddel says. "I was in on it early, but when I became director of the department, I thought I'd better back off." Instead, he unwound on weekends by rock climbing or by fly fishing with Boyer. Still, he did nothing to discourage the shenanigans by others. "In later years, when anything happened in the molecular biology department, I got blamed, even though I had nothing to do with it."

In the hallway, they played "bowling for dollars." The game involves rolling up a dollar bill tightly, and then tossing it just right, so that it comes to rest at a pre-decided line. Winners pocketed the dollars. "We'd usually get the business development and marketing guys when they came out in their suits," Goeddel recollects, but the 'suits' were not aware of the cahoots among lab personnel. "We had all these tricks where we could get the dollar bill to stop where it's supposed to." The secret? Partly, "you don't want a crisp, new dollar bill—you want a soft, old one," Goeddel says, giving nothing else away.

The times, and the people, were crisp and new, with plenty of 14-hour days of serious work. Staffers needed the tomfoolery to keep them sane. Because of the results of their research, "we knew there was a lot possible, but as scientists we had no understanding of what it would take to build a company like that. Those first five years, from 1978 to 1983, were the most fun I've ever had."

Goeddel went on to co-found South San Francisco–based Tularik with Robert Tjian

and Steve McKnight in 1991, and served as president and CEO until 2004, when the company was sold to Amgen, of Thousand Oaks, California, for $1.3 billion. He's now a venture capitalist with The Column Group in San Francisco.

Dennis Kleid

Dennis Kleid gave up a job at Stanford Research Institute (now SRI) to join Genentech, but he wouldn't come without his newly arrived post-doc, Dave Goeddel, so it was a two-fer for Genentech. The pair brought DNA synthesis capability to the company (Kleid had done a post-doc with MIT's Har Gorbind Khorana; Goeddel, with the University of Colorado in Boulder's Marvin Caruthers), believed to be the secret of their early success. Their competitors were trying to clone cDNA long before there were kits to do it, or even many restriction enzymes. 'Genentechers' cloned synthetic DNA—no problem with getting the clone in the proper register. They made it to order.

After the successful cloning of somatostatin, a proof of principle that the fledgling company needed to get buy-in from investors, Kleid was part of the team (along with Heyneker, Goeddel and City of Hope researchers Arthur Riggs and Keiichi Itakura) that cloned synthetic insulin. If somatostatin brought them fortune (or at least enough to keep them in business), insulin brought them fame. At the now infamous press conference at City of Hope to announce the cloning of insulin, Kleid found himself in front of hundreds of reporters. The first words out of his mouth were, "To say we cloned insulin is a lie," by which, of course, he meant that they had cloned the A and B chains, which, until they were assembled (outside the bug), were not insulin. Somebody managed to get the mike away from him (to avoid any further confusion), around the time the banks of lights that had been hauled into a rather small briefing room went out. Circuits blown, end of press conference.

At Genentech still (as of this writing), in 1986. Kleid changed hats to work with the legal team. "As the company grew, research side became less attractive, the mission became more focused," he says. In his present job, his longevity is useful because the company is still involved with litigation that goes back to the 1980s. He gets satisfaction from knowing that the company actually helps people. "I feel like I am Genentech. Seeing pictures of patients on the wall, it brings you down to earth."

Axel Ullrich

Axel Ullrich was part of the cloning team, along with Peter Seeburg, that came over from Howard Goodman's lab at UCSF. Though synthetic insulin was cloned first, the yields were not commercial, and the Lilly recombinant insulin product, Humulin, was made from his cDNA clone.

As a testament to the naiveté of these young scientists, not long after he joined, Ullrich decided to buy a car and sold some of his stock options—Swanson's brainchild to give the scientists a sense of ownership—for a few thousand dollars to buy a VW. After the company went public, the shares he sold were worth over a million dollars. Thereafter, it was called his million dollar VW (ref. 2).

Of the early group, Ullrich was perhaps the most concerned about maintaining academic creds—he had passed on a good job at Max Planck to join Genentech—and worked on basic research projects that had potential but not immediate application. "Genentech went after proteins with magic names [like tumor necrosis factor, interferon]. I followed my instincts, I wanted to work on receptors."

Spurred on by the discovery that human erb-2 had homology to a chicken oncogene, Ullrich was driven by the desire to find human oncogenes. He had the idea of targeting erb-2 with an antibody, but the Genentech management didn't see what antibody binding to a receptor would do therapeutically. In addition, at that time, companies were running away from antibodies because of a high-profile failure of an antibody drug for sepsis by the biotech company Centocor.

So, eventually, he left Genentech, but he continues his pursuit of targeted therapies to this day. Ullrich took a job at Max Planck Institute for Biochemistry, Martinsried, where he's been for 20 years. He says he has no regrets; his work has been well supported, and he has formed four companies on the side—two of which were bought (Sugen by Pharmacia, now Pfizer; U3 Pharma (for Ullrich 3) by Daiichi Sankyo), one that failed (for which he blames shortsightedness of venture capitalists) and one service company. And he has three drugs to his credit—insulin, the multi-kinase inhibitor Sutent (sunitinib) and Herceptin (trastuzumab), credit for which has not always gone to Ullrich.

After he left, Genentech picked up the erb-2 project and, after a few fits and starts, out came Herceptin. The rest is history… or maybe not. A made-for-TV movie on the development of Herceptin credits a UC Los Angeles oncologist with its development. Genentech isn't doing much better, according to Ullrich. A new documentary that Genentech commissioned got it all wrong, too, he claims. "It's amazing what Genentech can do with its power, its money."

Of the early days, Ullrich says, "I was very, very skeptical that it would work commercially, but if it did, I wanted to be a part of it."

1. Boyer, H. *Recombinant DNA technology at UCSF and the commercial application at Genentech*; typescript of an oral history conducted in 1994 in The UCSF Oral History Program and the Program in the History of the Biological Sciences and Biotechnology (Bancroft Library, University of California, Berkeley, 2001).
2. Ullrich, A. *Molecular biologist at UCSF and Genentech*; typescript of an oral history conducted in 2003 in The UCSF Oral History Program and the Program in the History of the Biological Sciences and Biotechnology (Bancroft Library, University of California, Berkeley, 2005).

Wasting cash—the decline of the British biotech sector

Graham Smith, Muhammad Safwan Akram, Keith Redpath & William Bains

Undercapitalization and overgenerous boardroom compensation for management have been major contributors to the poor performance of UK biotech.

Despite historic leadership in European biotech, the UK's industry has suffered a near collapse in the past two years and now has little private or public investment and no candidates for world-class companies. In this article, we argue that a combination of severe undercapitalization of UK companies, overgenerous boardroom cash remuneration, and lack of share- or performance-related reward for top executives has resulted in many executives not only spending company cash on nonproductive activities, but also in some cases frittering away already dwindling company financial resources in a manner that hastens the demise of their firms. As expected from conventional management theory, companies that are well capitalized, and whose senior executives have substantial shareholdings and reasonable cash reward have brighter prospects. Our analysis focuses primarily on public UK companies, but preliminary data suggest that private venture capital (VC)-funded companies have also decoupled CEO performance from CEO reward.

The failure of UK biotech

The UK has been seen as one of biotech's leading regions, second at a national level only to the United States. With ~435 biotech companies plus service providers and technology service providers, the sector employs around 18,900 people and generates revenues of around £2.5 billion[1]. UK companies account for 40% of biotech products in the pipeline worldwide (and 45% of those in phase 3)[1,2]. UK public ('listed') biotech firms have been a major source of the UK life sciences innovation with over £400 million spent annually on R&D.

Even so, growth of genuinely world-class companies in the UK has been elusive. Unlike the United States, British biotech has not had a well-known 'trophy' success, such as Amgen (Thousand Oaks, CA, USA), Genentech (S. San Francisco, CA, USA) or Genzyme (Cambridge, MA, USA). What's more, no launched blockbuster biopharmaceutical has originated in a VC-funded UK biotech—although Humira (adalimumab) originated from Cambridge Antibody Technology (CAT; Cambridge, UK), now Medimmune Cambridge, the company was funded by Peptech (Sydney, Australia), not VC; Campath (alemtuzumab), the other UK-originated biotherapeutic blockbuster, originated in an academic lab.

Stock values of the UK biotech sector have tumbled in the past year reflecting this lack of delivery. The performance of UK small caps in particular was very poor in 2007, with biotechs in the UK and continental Europe registering an overall fall of close to 37%. For UK biotech companies, unlike their European peers, the 2007 decline followed a poor year in 2006 as well[3]. This has dampened interest in the sector and has made equity fund-raising very difficult. In addition, the VC market has had poor returns from UK biotech investments

Kino Brod © Images.com/Corbis

Has the greed of UK management accelerated the decline of the country's biotech sector?

Graham Smith is at East Grinstead, West Sussex, UK; Muhammad Safwan Akram is at the Institute of Biotechnology, Department of Chemical Engineering and Biotechnology, Tennis Court Road, Cambridge, UK; Keith Redpath is at JM Finn Capital Markets Ltd., London, UK; and William Bains is at the Institute of Biotechnology, Department of Chemical Engineering and Biotechnology as well as Rufus Scientific Ltd., Melbourn, Royston, Herts, UK.
e-mail: william@rufus-scientific.com

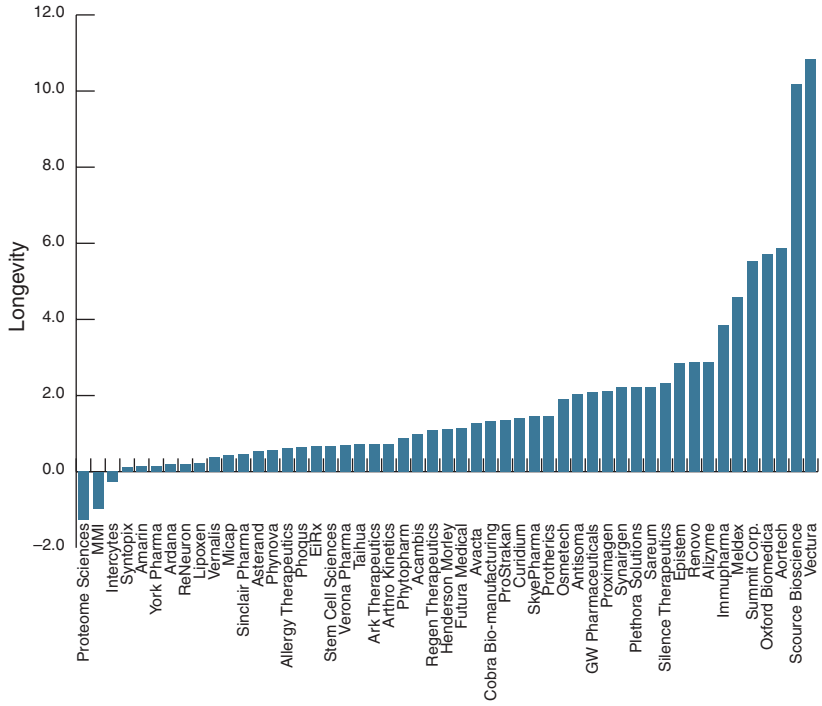


Figure 1 Longevity analysis of UK biotech companies. Longevity was calculated as the amount of cash reserves in the company divided by their average annual cash spend. *X* axis: rank ordered according to company longevity. Both values were obtained from public accounts in year or quarter to June 2008 for all biotech companies listed on a London stock exchange. Companies with negative longevity have net debt on their books.

and consequently many funds have withdrawn from biotech investing in favor of other sectors. We and other industry observers see no appetite for biotech stocks on the London markets; although indices such as the TechMark index on the London Stock Exchange (LSE) fell 'only'18% in 2007–2008 (before the current banking crisis), liquidity and the ability to raise more cash have fallen further.

At the start of 2008, several news articles strongly advised against investing in biotech stocks, focusing on several well-known failures. An article in the *UK Financial Times* stated, "Investors should stay clear of the UK biotech sector and instead put their capital in alternative low-risk industries, as biotech is unlikely to generate positive returns this year"[4]. The article quoted Paul Cuddon, an analyst at the London-based investment bank KBC Peel Hunt, as saying, "Profitability is a luxury for UK biotech and would in our opinion represent a misuse of capital for many companies. Not only have share prices fallen but liquidity has also dried up"[4]. Analysts from such investment houses as Seymour Pearce (London), Edison (London) and Numis (London) have issued surveys of the biotech sector, citing failure to perform as a key factor in falling prices[5].

A longevity analysis performed in mid-2008 reveals 67% of companies have less than two years of cash left and 45% have less than one year's cash left (**Fig. 1**). For all small companies, capitalization is the single biggest factor in success or failure, ahead of product, management skills or external economic environment[6]. The companies with a high probability of long-term growth have partnerships with big pharma, large cash reserves, are profitable, have strong pipelines or products about to launch. Those with little cash reserves face a bleak future. If a company's technology does not deliver product candidates, or those candidates fail regulatory hurdles, a biotech with cash has the option of using its cash reserves and scientific expertise to acquire new technology, through in-licensing or acquisition of a cash-strapped peer or one on the verge of bankruptcy. If it has little cash, it cannot do this and must pursue what it has in its portfolio. The poor longevities in **Figure 1** illustrate why many UK companies have few options available.

Why has UK biotech failed so spectacularly? The macro-economic climate is in part to blame, and it might seem contrary at this moment to look beyond the spectacular failure of the global banking system and its managers for the cause of UK biotech's downfall. Even so, we maintain that the failure of UK biotech predates the 'credit crunch' and, elsewhere in the world, the biotech sector has not fared so badly. Indeed, UK biotech investment trusts performed well in 2008 (ref. 7). As is the case elsewhere in the world, the inherent riskiness of the biotech business model is a factor in tempering investor enthusiasm. The chances of a small company with a handful of early drug development programs reaching a major commercial success is small because of the vagaries of drug discovery and development[8]. A run of failures in specific companies—and the UK sector has seen its fair share—dampens enthusiasm for biotech as a whole, even though an objective observer should realize that one company's clinical trial failure does not affect the chance of success of other companies in unrelated fields. So informed investors should be able to take a view of the risks involved and, if they wish to support the sector at all, support it with sufficient cash to have a chance of success. Either UK biotechs are not getting enough cash to execute their businesses, or they are not using that cash effectively. Our analysis suggests that failure is due to both effects, and specifically to two interacting factors: first, undercapitalization, especially in the early stages of a company's evolution, resulting in weak companies being brought to the public markets; and second, excessive expenditure on items other than R&D, and especially on board- and executive-level salaries. Together, this lethal mix has resulted in an industry that cannot perform, and so is not supported.

Undercapitalization

The undercapitalization of the private UK biotech industry is now well established. UK companies receive between 10% and 30% of the investment of their US counterparts[9–12]. Funding levels in the United States gradually rose in the 1980s and 1990s as the industry matured and investors saw the value in building world-class companies that would attract world-class management. In contrast, funding levels fell in the UK over the same period[13] and plummeted in 2008 (**Fig. 2**), substantially exacerbating companies' financing plight. The stark contrast between the United States and the UK is exemplified by one of the UK's major biotech hubs in Oxford and Cambridge. This region has been home to several companies (e.g., CAT, Celltech, KuDoS, Domantis and Acambis) that brought in some of the highest value trade sales in UK biotech over the past five years. But most of these companies were founded over a decade ago and the financing to create another CAT or Celltech has not been available for years; the largest company left independent in Cambridge is Xenova, whose components were founded before 1990.

Underfunding of this sort results in companies being valued poorly when they float on the stock market, raising only small amounts of cash from the sale of shares at initial

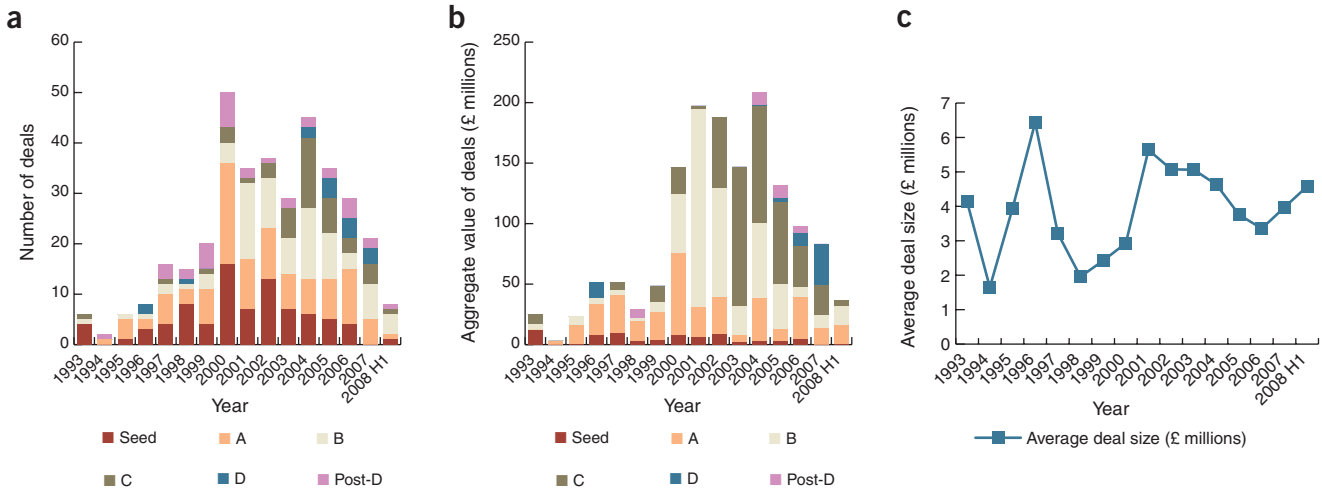


Figure 2 Private investment in Oxford and Cambridge biotech clusters. Publicly announced investment in private biotech companies in the Oxford and Cambridge clusters, 1993–2008. (**a**) Number of deals. (**b**) Aggregate value of deals. (**c**) Average size of deal for private (business angel and institutional) investors, not including corporate investment. Data from company web sites, Capital IQ (New York, NY, USA) and the Eastern Region Biotechnology Initiative (Babraham, Cambridge, UK). Grant income is not included in these figures. Deals are separated by stage (seed, A, B, C, D and others) from private or institutional investors and corporate equity investment.

public offerings (IPOs). In line with other studies[9,14,15], our data show that small levels of pre-IPO investment result in low-cap public companies (**Fig. 3**). Notably, several recent UK trade sales have valued the companies at below the total invested in them. Several of these sales have been of companies that have had a great deal of technical success, so the high failure rates in therapeutics discovery and development are not the sole cause of the losses being made.

This would be a surmountable problem if companies invested their limited resources in product development. High-tech, cutting edge therapeutics of the sort in which biotech companies should excel[8] are out of the grasp of micro-cap companies, but they could focus on a no-research, development-only (NRDO)[16], low-risk but niche market, sales-driven or other type of business model that is commensurate with their cash reserves (if their shareholders let them[17]). It appears, however, that they do not do so.

Of particular note, our analysis reveals that UK biotech companies spend a disproportionate amount of money on remuneration for top executives and directors, and it is this drain on resources that has become a critical factor in UK biotech companies' lack of performance as cash investment has dwindled in the last eighteen months.

Boardroom salaries

We analyzed the cash position and boardroom salaries for all 51 public UK biotech companies. Data collected included all forms of remuneration for all executive and nonexecutive directors at all UK biotech companies listed on a London exchange in July 2008 (**Box 1**). Data were collected from the last published annual report and press releases, and so provided a picture both of the company's audited cash position and of the salaries at the time of that cash position, enabling a direct analysis of how salaries would affect future cash.

Boardroom remunerations drain the cash reserves of UK-listed biotech companies at a total rate of at least £40 million per year. Sixty-five percent of UK CEOs are paid an annual remuneration of over £200,000, with the average CEO pulling in a remuneration just short of £300,000 at £297,947. (Salary remuneration figures include all cash or cash equivalent remuneration, including bonuses and pensions, but not equity-related compensation.) At the top end of the scale, the top half of biotech CEOs took home an average of £443,793 which is equivalent to the top 10% of CEO salaries in the United States (US CEO national averages found at http://www.salary.com/ on May 8, 2008). The top UK biotech CEO's remuneration of £813,000 was 10% higher than the average Financial Times Stock Exchange (FTSE)100 CEO's salary of £737,000 (ref. 18). There is

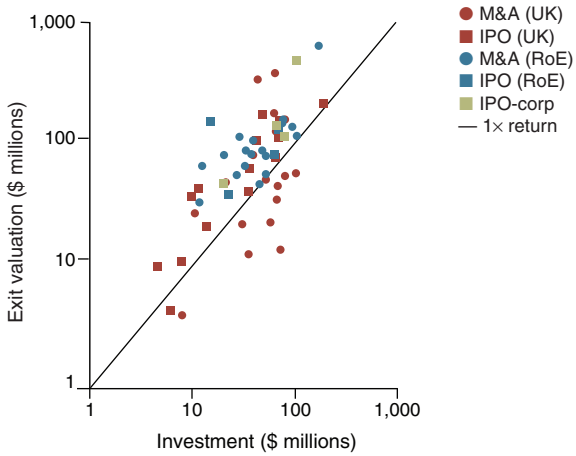


Figure 3 Private finance versus exit valuation. Amount invested versus exit valuation for 59 European biotech companies that had received VC investment and had exited in 2001–2007 inclusive. Exits are categorized into IPO (flotation on a public market) or merger and acquisition (any other type of exit, whether called merger, sale, acquisition or other form). Companies are divided into startups from UK, startups from the rest of Europe (RoE) and companies formed in UK or the rest of Europe with a substantial legacy of intellectual property, technology or products from a parent corporation, for whom the invested cash is not necessarily a true reflection of invested value they received ('corp'). Source: VentureSource, Capital IQ, company web sites and press releases and reference[25].

Box 1 Companies surveyed

Companies considered in this survey all had a primary listing on a UK stock exchange. They also all had a life science focus (as defined by UK listing categories as 'any technological application that uses biological systems, living organisms, or derivatives thereof, to make or modify products or processes for specific use') and/or provided R&D services or support. Companies that are 'pure' pharmaceutical companies (e.g., Shire, Astrazeneca and GlaxoSmithKline) or that provide services to the healthcare provision industry and consultancies were all excluded. UK company accounts run for a 12-month period starting at an arbitrary date (set by the company), so annual reports analyzed in July 2008 included data from accounting periods of 12 months from mid-2006 onwards, depending on company policy. Most companies chose to start their accounting year on January 1 or April 6.

minimal correlation between the longevity of a biotech or its productivity in generating new products and the salary of the boardroom (**Fig. 4**; correlation coefficient $r^2 = 0.0134$ for correlation of boardroom remuneration and longevity, $r^2 = 0.0915$ for correlation of boardroom remuneration to pipeline), which illustrates that a more expensive management team doesn't result in improved business or a better long-term financial position.

The majority of board-level compensation in this group of public companies is in cash or cash equivalents. The directors of UK biotechs generally have a very low interest in the companies they are managing with the median board ownership a mere 4.7% of the company they direct (**Fig. 5**). In contrast, **Figure 5** also shows the shareholding of all the software/semiconductor ('tech') companies with a market cap of less than £250 million listed on LSE and London's Alternative Investment Market (AIM): the median board shareholding is 16.5%. Biotech CEOs of public companies are often not the company founders and have an especially small interest in the firm, with the median CEO holding just 0.55%.

Most UK public companies do not retain founders on their boards. In a previous study[19], one of us (W.B.) showed that companies that do retain founders perform better, which is in line with the expectation that management and directors motivated by significant shareholdings will perform to maximize shareholder value (although due to their depressed valuation, stock options in European public biotech companies might no longer be such a powerful motivator[20]). Our informal analysis of the personal history of the board members of the companies analyzed here confirms that companies with more robust cash positions tend to have founders retained as executives and board members[19]. Of the top-earning biotech CEOs (£400,000+), the median share holding is only 0.42%.

What is more startling is that as a biotech approaches bankruptcy, the boardroom and CEOs appear unwilling to sacrifice their large salaries for the welfare of their company. Six out of ten of the biotech firms with the shortest longevity had CEOs with salaries above £200,000, two CEOs even had salaries of over £500,000 when their companies were on the verge of bankruptcy. One company at the time of their annual report had less than 2 months worth of cash left, but the management board in total had a remuneration of £925,000, £243,000 above the average for the sector. If the board were to halve their wages and take share options instead, this company could have increased its life expectancy 250%. A second company was undergoing a finance shortage in late 2006: it had to reduce its R&D spending from 2006 to 2007. In the same period, however, the boardroom salaries increased by 25%, against the trend of spending and draining the cash reserves of the company even further. In the same period, the board of directors had also reduced their shareholding from 12.3% to 9.55%, slowly reducing their commitment to the company. These two examples illustrate a growing trend of biotech directors to slowly diminish long-term commitment and increase immediate cash rewards, a trend also noted anecdotally by Rychlik and DiPierro[21].

The principle issue here is not whether these CEO salaries are higher than those in other sectors or other geographies, but that they appear disproportionate to the performance, and particularly to the cash reserves of the companies concerned. But are they also higher than 'typical' CEO salaries? Comparable data

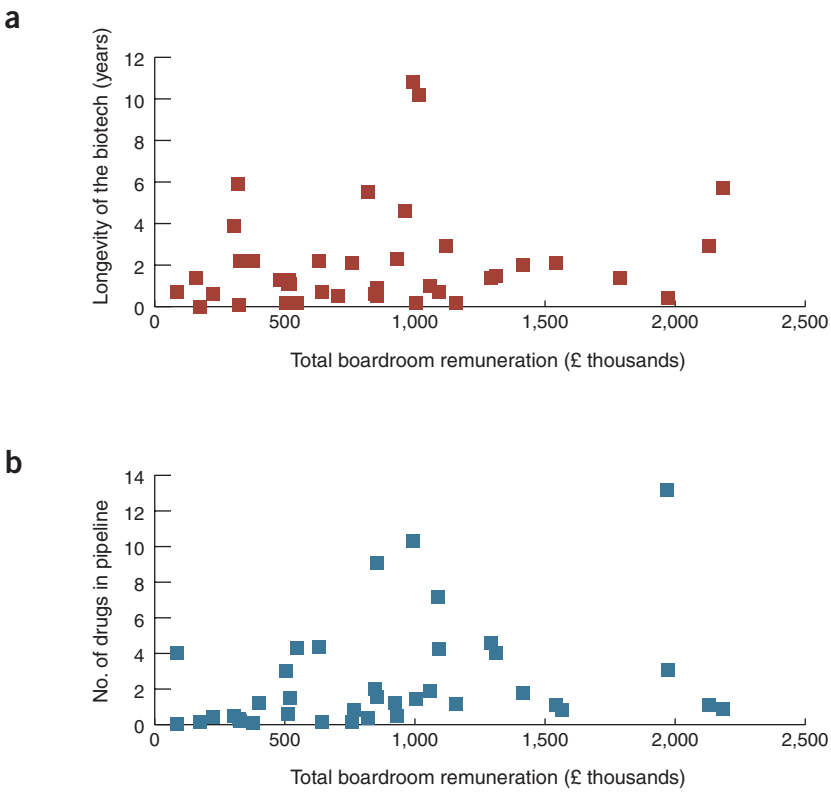


Figure 4 Boardroom remuneration and company success. (**a**) Boardroom remuneration and company longevity. *X* axis, total board remuneration (including salary, bonus and pension but excluding equity-based reward); *y* axis, company 'longevity' (cash reserves/annual burn rate). (**b**) Boardroom remuneration and drug product pipeline. *X* axis, boardroom remuneration (as above); *y* axis, drug pipeline (that is, the number of products on the market (launched) plus the number expected to be launched from the company's pipeline of development programs and calculated as: number of launched products + [(number in registration) × 0.8] + [(number in phase 3) × 0.37] + [(number in phase 2) × 0.2] + [(number in Phase 1) × 0.08] + [(number in preclinical) × 0.03]).

for other European companies are not available, as other European markets do not have the disclosure requirements of London stock exchanges. Even so, many European IPO prospectuses declare CEO and board salaries. Data from 32 European IPO prospectuses that disclose CEO or board cash remuneration give an average declared CEO cash remuneration of ~£220,000, and average board-level remuneration of ~£460,000, slightly lower than the current public UK averages. (Values reported in local currency or Euro were converted to pounds sterling at the exchange rate pertaining at the time of the IPO.)

It isn't only the salaries of biotech CEOs that appear inflated—matching up with those of executives in FTSE 100 companies—golden parachutes for directors of UK biotech companies are also often on a par with much larger, profitable companies. For example, at one leading UK biotech in our study sample, the CEO resigned from his position 30 days before the annual report was filed but received £491,000 in that time for compensation for loss of office. This made his year's salary, bonuses and pensions total >£1 million. As a comparison, Lord Terence Burns, the chairman of one of the UK's largest retailers, Marks & Spencer's, received a parachute of £450,000 when he left in March 2008. Marks and Spencer is a FTSE 100 company with annual revenues of £8.6 billion—revenues 7,500 larger than those of the biotech company in question.

Private biotech board salaries

There are far fewer data available about the salaries of board members in private biotech companies in the UK. Such salaries are not required to be declared separately in formal accounts, and so are confidential information of the companies concerned. At a semi-anecdotal level, we have collected data on the cash remuneration of 28 private UK biotech CEOs, and also asked the respondents (who were investors or independent nonexecutive directors), under guarantee of absolute anonymity, to evaluate whether the CEO was doing a good job (**Fig. 6**). Those providing the data were also asked whether the CEO had been successful in achieving the critical goals for the company at that time (that is, whether they had achieved the goals that the investors considered critical for company survival, growth and development, usually fundraising and/or closing deals, and hence whether the investors themselves considered that the funds spent on their salaries was deployed effectively). Although angel- and revenue-funded companies follow economic theory by rewarding CEOs for success, UK VC-funded companies appear not to compensate CEOs on the basis of performance, similar to the situation noted above in UK public biotech companies (**Fig. 4**). Given that the average investment round is now around £4 million (**Fig. 2**), which is meant to last a company between 1.5 and 3 years, this suggests that CEOs could be a significant, nonproductive drain on some private companies' performance. Survey data from the United States suggest that US private biotech CEOs are paid much more, around $380,000 cash remuneration in 2008 (around £200,000) at the exchange rates prevailing when this study was carried out[21]. This, however, should be put in the context of the substantially greater funding that US companies receive, as noted above.

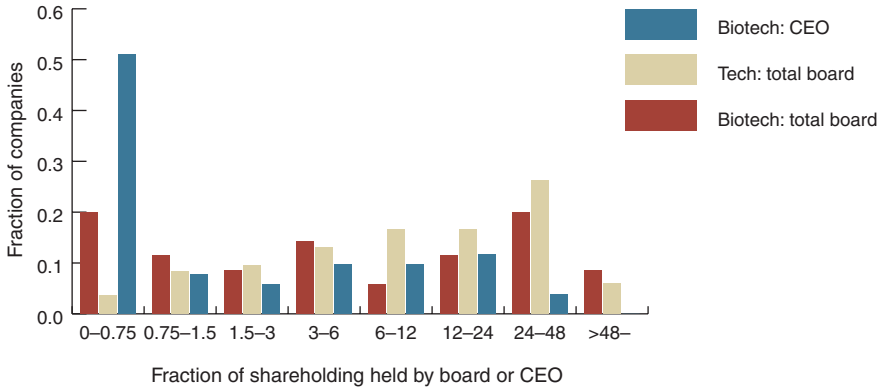


Figure 5 Boardroom shareholding. Fraction of issued shares held by CEO or by all the directors of public UK biotech companies, and fraction of issued shares held by all directors of mid- and low-cap tech (non-biotech, science or technology-based) companies. Data are shown for a total of 35 biotech companies, 85 tech companies and 51 biotech CEOs. Source: company reports and Thomson One Banker Regulatory News Service filings.

DISCUSSION

Why do shareholders allow UK public biotech companies to accumulate top management that pays itself so much, is unmotivated to drive shareholder value and as a consequence apparently drains the company of resources, notably cash?

In part this is an unintended effect of changes in the law requiring publication of public company CEO salaries: with the salaries of their peers in more developed companies as a benchmark, CEOs can demand at least the same salary, and often a higher one to compensate for the higher risk in their position running a small, cash-strapped enterprise. In part, it is also a consequence of the high salaries in VC-backed private companies: if a failing CEO in a private company can be paid £150,000, why should an equivalently successful one in a public company with twice the market cap not be paid twice as much? But these are excuses, not reasons.

One reason cited is the lack of high-quality commercial management in UK biotech, which means that market demand drives up the price of the few high-quality CEOs. In our opinion, this is a fallacy. There is no such thing as 'UK management'. Biotech management is at least European, and in many respects global. A visit to Boston will show that many of the executives in biotech companies there are ex-patriot Europeans. A survey of European private VC-backed biotech companies in 2006 showed that there is a substantial number with CXO (chief executive, scientific, financial, operating

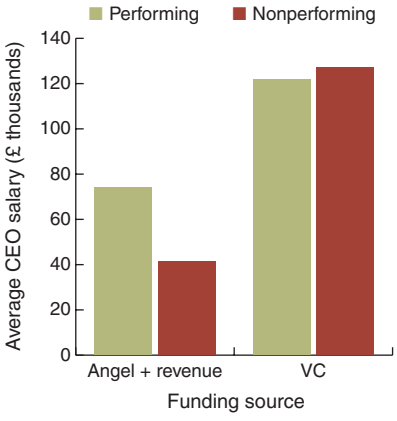


Figure 6 Private company CEO salaries. Reported salaries of 28 UK private biotech company CEOs (data for years 2001–2008). Companies were categorized by respondents into those funded by angel funding, revenue or both, and those funded by institutional investment (VC). CEOs were categorized into performing (achieved well given the business, environment, cash position) or nonperforming according to the judgment of the reporter. Everyone polled had direct personal knowledge of the CEO salary and performance. Salary is cash salary, cash bonus and cash equivalent benefits such as pension. Equity-based compensation, such as shares or options, is not included.

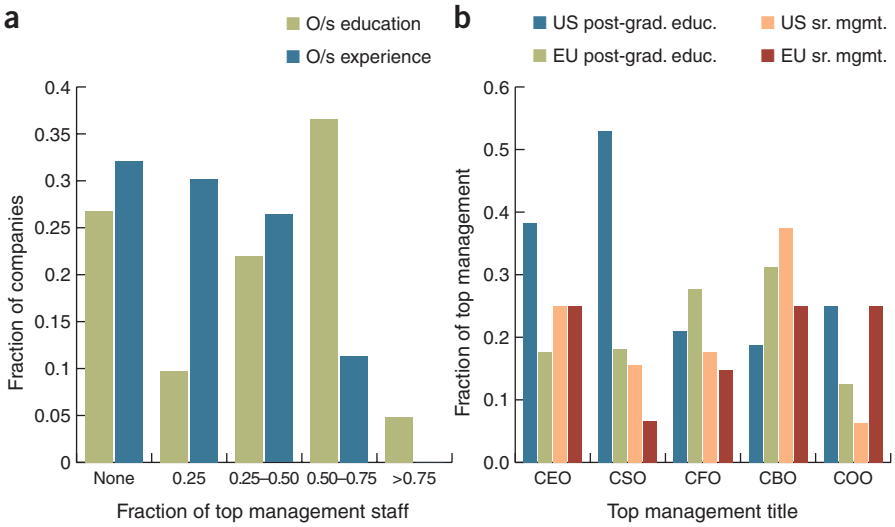


Figure 7 International experience in European biotech CXOs. (**a**) Fraction of CXO staff in 90 European VC-backed biotech companies that have substantial postgraduate education (PhD, postdoc, MBA or other equivalent professional qualification) or significant business experience in top management roles outside the country in which the company is based. Data were collected from company websites and Capital IQ (http://www.capitalIQ.com/) in 2007. (**b**) Fraction of CXOs who have had post-graduate educational or senior level management experience outside the country in which the company is based, divided into experience in other EU countries or experience in the United States. Graphs reproduced from ref. 26 with kind permission of Palgrave Macmillan. CXO, chief executive level offices, which can include chief executive, scientific, financial, operating or business operators. O/s, overseas.

or business officer) level executives from other countries or with international experience (**Fig. 7**), with up to 50% having senior level international experience for business development and CEO positions. The management market is mobile and global. Either there is a global shortage of commercially competent CEOs or there is no shortage and we must look elsewhere for the reason for high boardroom remuneration in UK biotech.

Why do shareholders allow companies to get away with such salary structures in private or public companies? We have identified three potential explanations.

First, the companies are inherently such unattractive commercial propositions that enormous salaries have to be offered to executives to join them and stay with them, as otherwise the international market for executives would mean that the executives go elsewhere. A clear source of company weakness is undercapitalization, as noted, but high board pay also exacerbates the undercapitalization problems of UK companies, creating a vicious circle of decline. Second, shareholders do not actually care or do not recognize what success and failure are until too late. A final explanation is that shareholders do recognize there is a problem, and the businesses are potentially attractive, but cannot act in isolation: it needs all investors and all executives to act in concert, as otherwise the few companies offering realistic remuneration packages will not attract top management.

The first thesis suggests a gloomy outlook for UK biotech indeed. But it is supported by the observations of genuinely successful entrepreneurs that outstanding management follows outstanding opportunities. Management is willing to be motivated by success[13]: by implication, poor opportunities will attract mediocre management that will not wish to be rewarded on the basis of success because they are not expecting it. The only solution is to create, support and fund better companies.

The second thesis seems quite implausible, but we note that there are very few activist shareholders in UK biotech public markets. It is also possible that those few that do follow biotech themselves have such enormous remuneration packages that biotech CEOs appear quite underpaid by comparison (e.g., see refs. 22,23: our informal understanding is that at least some specialist, private investor partners have comparably sized remuneration packages). This, however, is a failure to understand the business in which they are investing and also does not bode well for the sector.

The third thesis is superficially an attractive explanation. If an investor in a company tries to force the company to reduce boardroom remuneration, surely the good management will leave to join other, more munificent companies. But our data do not suggest that this is the case. Some high-performing companies with high boardroom shareholding (**Fig. 5**) pay their CEOs relatively modestly. In line with this, there is a noticeable trend for boards with higher shareholding to pay themselves less, and a similar but less noticeable trend for CEOs with higher equity participation to receive less cash reward (**Fig. 8**). The reward of sharing in the success of the business apparently outweighs the lower cash rewards today. If an investor were to force down salaries, increase board equity participation and adequately finance the company for success, the result should be a company that is attractive to commercially orientated management. Undercapitalization and overspending on management salaries are linked issues.

Boardroom salaries are not the sole cause of the UK biotech sector's woes. A complex corporate tax structure that is among the highest in Europe, a tightening regulatory environment, and, of course, public market and financing problems play their part. Even a domestic housing market that could (until the start of 2008) guarantee a tax-free 10% return on investment

Figure 8 Boardroom shareholding versus salary. Cash remuneration (salary plus bonus) is plotted on *x* axis against shareholding (*y* axis) for CEO and all board members (percent of shares held by CEO or board). Data from same source as **Figure 1**.

for merely owning a big property in appropriate parts of Oxford, Cambridge or London affects the biotech sector, by draining capital from productive investment into the task of merely existing. But even if these problems are solved, the continued drain of excessive, performance-unrelated boardroom salaries that are unconnected to company success will continue to destroy the UK's already battered biotech sector, and hinder its recovery when the market turnaround that informed observers expect happens in 2011 or 2012 (ref. 24).

COMPETING INTERESTS STATEMENT
The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturebiotechnology/.

ACKNOWLEDGMENTS
This article is based on theses submitted by G.S. and M.S.A. in partial fulfillment of the Cambridge University Masters in Bioscience Enterprise programme (http://www.biot.cam.ac.uk/mbe). The authors are grateful to JM Finn and Company and the Eastern Region Biotechnology Initiative for support during the preparation of these theses. The authors emphasize that specific mention of a company in this paper is not meant to imply that the company is 'better' or 'worse' than average.

1. UK Trade and Investment. *Bioscience UK* (UKTI, London, 2007). <https://www.uktradeinvest.gov.uk/ukti/fileDownload/UK_Bio_Science_reflow_2_AW.pdf?cid=406181>
2. Ernst & Young. Beyond borders: a global perspective. Biotechnology industry report 2006. (Ernst and Young, New York, 2006).
3. Hemming, R. *Has British biotech bottomed? Investors Chronicle (London)*, January 24, 2008. <http://www.investorschronicle.co.uk/Companies/ByEvent/Risk/Analysis/article/20080124/59fed530-c9b3-11dc-8b38-00144f2af8e8/Has-British-biotech-bottomed.jsp>
4. Davoudi, S. Biotech sees market value slip away. *Financial Times (London)*, January 17, 2008. <http://www.ft.com/cms/s/0/28a9396e-c466-11dc-a474-0000779fd2ac.html>
5. Davoudi, S. *Financial Times (London)*, 2008. <http://www.ft.com/cms/s/0/51ffe382-bf1e-11dc-8c61–0000779fd2ac.html>
6. Riquelmo, H. & Watson, J. *Int. Small Bus. J.* **20**, 395–420 (2002).
7. Walters, L. Biotech investment trusts looking good despite challenges. Trustnet News and Research, Jan. 27, 2009. <http://www.trustnet.com/News/Research.aspx?id=34596&researchCode=Fund%20selection>
8. Bains, W. *Drug Disc. WORLD* 2004, **9**–18 (2004).
9. Bains, W. J. *Commercial Biotechnol.* **12**, 274–283 (2006).
10. Critical, I. Ltd. The equity gap: European and US biotechnology 2001–2 (Bioindustry Association, London, 2004).
11. Mitchell, P. *Nat. Biotechnol.* **23**, 1029 (2005).
12. Critical, I. Ltd. Comparative statistics for the UK, European and US biotechnololgy sectors: analysis year 2003 (Department of Trade and Industry, London; 2005). <http://www.berr.gov.uk/files/file28439.pdf>
13. Klausner, A. *Bio/Technology* **5**, 1129 (1987).
14. McGully, M.G. in GTCbio/s Metabolic Diseases World Summit: Partnering and Deal-making Summit (Burlingame, San Francisco, CA, USA, (2005).
15. Edelson, S. *Biocentury* **14**, 1, A1–A6 (2006).
16. Thiel, K.A. *Nat. Biotechnol.* **22**, 1087–1092 (2004).
17. Maynard, O. & Bains, W. J. *Commer. Law Stud.* **81**, 1–37 (2008).
18. Vorster, G. FTSE-100 bosses take home 100 times more than average employees. *Personneltoday.com*, October 29, 2007. <http://www.personneltoday.com/articles/2007/10/29/43022/ftse-100-bosses-take-home-100-times-more-than-average.html>
19. Bains, W. J. *Commercial Biotechnol.* **13**, 139–149 (2007).
20. Lawrence, S. *Nat. Biotechnol.* **23**, 157 (2003).
21. Rychlik, B. & DiPierro, M. *Nat. Biotechnol.* **27**, 98–99 (2009).
22. Bawden, T. Lehman and Bear Stearns join banking bonus boom. *Times (London)* Dec. 15, 2006.
23. Chong, L. Goldman staff share $16bn in bonus bonanza. *Times (Lond.)* Dec. 12, 2006.
24. Bate, L. *The Current State of Biotech in Europe.* MPhil thesis. Univ. Cambridge (2007).
25. Akram, M.S. *Investment Trends in the Cambridge Biotechnology Cluster* (Institute of Biotechnology, University of Cambridge, Cambridge, 2007).
26. Bains, W. *Venture Capital and the European Biotechnology Industry* (Palgrave Macmillan, Basingstoke, UK, 2008).

A day late and a few million dollars short

Dianna Goldenson

The pitfalls of seeking and obtaining a patent term extension.

It's Monday, and your homework assignment is due "within two days." When do you have to hand it in—Wednesday? Or maybe Tuesday? Missing a deadline in school could be a problem but would not ordinarily cost money. In the world of patented drugs, however, missing a deadline could cost millions, even billions, of sales dollars.

At the US Patent and Trademark Office (USPTO), a tricky filing deadline is being frequently miscalculated by applicants for patent term extension (PTE), a procedure that can add up to five years to a patent's life. Until recently, the government was miscalculating this deadline as well, resulting in the grant of several invalid extensions based on late-filed applications. The law is not the problem because the filing period is clearly defined in the statute. The problem is that people are simply counting wrong. Missing the PTE filing deadline is an incurable error, but missing it without getting caught opens a whole other can of worms.

Table 1 The last 100 applications filed and granted PTE and viewable on the USPTO's online PAIR database

PTE filing year	No. of applications with miscalculated deadline	No. of applications with correctly calculated deadline
2002	18	6
2003	17	9
2004	15	0
2005	17	4
2006	11	2
2007	0	1
Total	78	22

Hatch-Waxman and the value of long-lasting patents

In 2008 alone, total worldwide revenues for the pharmaceutical industry included $63.7 billion for Johnson & Johnson[1], $48.3 billion for Pfizer[2] and $31.6 billion for AstraZeneca[3]. Popular branded drugs can bring in several millions of dollars per day before generic competition enters the market. With so much at stake, particularly for blockbuster drugs, brand companies dedicate themselves to maintaining market exclusivity; and they need long-lasting patents on their drug products to do it.

A patent holder has the right to exclude others from making, using, selling or offering to sell the patented invention for a certain time period[4]. A patent on a drug becomes most valuable when the drug receives marketing approval from the US Food and Drug Administration (FDA). Because a drug must undergo several years of testing before the FDA approves it, the life of a patent protecting the drug is often ticking away, shortening the duration of patent protection that will be left when the drug reaches the market. To compensate drug developers for some of this lost time, a law was passed as part of the Hatch-Waxman Act that allows a patent to protect the drug beyond its original expiration date, for sometimes as long as five years[5].

Patents on branded drugs raise barriers that generic drug companies must overcome before they can market their generic versions. For instance, Hatch-Waxman requires brand pharmaceutical companies to list all their patents covering the FDA-approved product (patents that claim the active ingredient, approved formulation and/or approved method of use)[6] in an FDA publication commonly known as the "Orange Book"[7]. When a company files an Abbreviated New Drug Application (ANDA) seeking approval for a generic drug, it must make an official statement (referred to as a "certification") as to how the marketing of its generic drug will avoid infringing for each patent listed in the Orange Book for the branded drug[8]. Patents listed in the Orange Book can thus be substantial obstacles, especially if a listed patent was granted an extension and covers the active substance itself. Generic companies often need to wait for such patents to expire before marketing generic versions because it can be difficult to avoid infringement of a so-called "substance patent." The longer the extension, the longer the wait, unless the generic can prevail in costly and protracted patent litigation.

A tricky path To PTE

To obtain a PTE, a patent owner must file an application with the USPTO that satisfies several requirements under 35 USC §156. One of the more innocuous-looking requirements is that the application must be filed in a timely fashion. But many applicants have a hard time identifying the deadline. In fact, not only have many patent lawyers gotten it wrong, but the USPTO and FDA have made this error many times as well.

According to the PTE statute, "an application may only be submitted within the sixty-day period beginning on the date the product received permission under the provision of law under which the applicable regulatory review period occurred for commercial marketing or use"[9]. The statute seems clear enough: FDA approval triggers a 60-day period for filing a PTE application, starting with the approval

Dianna Goldenson is at Darby & Darby, New York, New York, USA. The views expressed in this article are those of the author and do not reflect those of Darby & Darby or any of its clients. e-mail: dgoldenson@darbylaw.com

date. In the world of federal laws and institutions, however, when a deadline is given, day one of the relevant time period is almost invariably the day after the trigger date. For example, most people would presume that a homework assignment that is given on a Monday and due "within two days" must be handed in by Wednesday. This is the common, well-entrenched convention[10]. The PTE statute conflicts with this convention by stating that the 60-day period begins on the date of regulatory approval. Thus, the first day of the filing period is the trigger date. For example, if FDA approval occurs June 15th, the PTE application filing deadline is August 13th. Correctly identifying this deadline is critical. If you miss it, the patent is ineligible for extension.

In the last 100 filed applications that have been granted PTE and are viewable on the USPTO's online PAIR database[11], 78 incorrectly identify the 60th day of the filing period[12], as shown in **Table 1**. The USPTO may have recognized the danger of this miscalculation because its own manual, the *Manual of Patent Examining Procedure* (MPEP), cautions PTE applicants to file early "to avoid the application being denied because the filing deadline was inadvertently missed"[13].

Near misses, casualties and possible legislative reform

Given the high stakes and draconian consequences, one would expect PTE applicants to file early, and most do. Much of the necessary information can be gathered even before FDA approval, so planning ahead is very helpful as 60 days can lapse quickly during the busy period just after a company receives FDA approval. Yet, this is undoubtedly a dangerous situation, particularly because the 60th day has been regularly miscalculated for over 20 years by the USPTO, the FDA and most patent lawyers who submit these applications. In dozens of cases, PTE applicants have unwittingly filed on the 60th day, evidently thinking that they were filing a day early. This group of lucky applicants includes, for example, the makers of Lovenox (enoxaparin) and Crestor (rosuvastatin calcium). The patent for Lovenox (at one time, the best selling anti-thrombotic drug in the world) was extended by three years, during which this product brought in an additional $5.58 billion in global sales. The patent on Crestor (a cholesterol medicine with $3.6 billion of global sales in 2008) will expire in 2016 after an extension of about 3.5 years. Without this bit of luck, these PTE applicants would have been a day late and a few billion dollars short.

Of course, not all applicants have been this lucky and a number of large drug companies have forfeited their PTE eligibility by missing the filing deadline (**Table 2**).

One case that achieved notoriety relates to the Angiomax (bivalirudin) PTE application[14], where the USPTO initially miscalculated the deadline[15], but later corrected itself, nevertheless concluding in both instances that the application was late[16]. In support of its final decision, the USPTO cited *Unimed, Inc. v. Quigg* as a court decision that addressed the issue of PTE application timeliness and stated that "section 156(d)(1) admits of no other meaning than that the sixty-day period begins on the FDA approval date"[17].

Prompted by the Angiomax case, lawmakers in 2006 tried to amend the PTE statute by proposing a bill dubbed the "Dog Ate My Homework Act"[18]. Under this bill, requests for an extension would be allowed a three-day grace period at the discretion of the USPTO director for unintentional delays, which would presumably include delays due to inadvertent miscalculation of the filing deadline[19]. This grace period would retroactively apply to requests for PTE that are pending before the USPTO or still subject to judicial review at the time of enactment. But this safety net does not come cheap. For PTE applications on anticoagulant drugs such as Angiomax, the bill would impose a late fee of $65 million (ref. 19). For other products, the late fee would depend on commercial success, that is, the sum of "any net increase in direct spending arising from the extension of the patent term," "any net decrease in revenues arising from such patent term extension," and "any indirect reduction in revenues associated with payment of the fee"[20]. The proposed bill was first included in the Patent Reform Act of 2007 (HR 5120), but efforts to pass this legislation stalled. The bill was resurrected in the Responsive Government Act of 2008 (HR 6344), which quickly passed the House of Representatives in June amid strong supporting statements from a few congressmen, one of whom went so far as to say that the bill "would save lives"[21]. The bill was referred to a Senate subcommittee last July but has not been passed as of this writing.

The "Dog Ate My Homework Act" could also benefit the makers of A180 (danofloxacin; an animal drug for treating bovine respiratory disease), Prilosec OTC (omeprazole; a human drug for treating heartburn) and Symbicort (budesonide and formoterol; a human drug for treating asthma), all of which were denied PTE for these products, at least partly owing to a late PTE application.

In the case of A180[22], the FDA first concluded (mistakenly) that the application was timely filed[23]. Later, the USPTO corrected the FDA's error and concluded that the application was untimely[24]. In the case of Prilosec OTC[25], the FDA again mistook the application as timely[26], after which the USPTO sent the FDA a letter correcting this mistake and stating that the PTE application was ineligible on grounds that included untimely filing[27]. About six months later, in October 2008, the FDA responded to the USPTO with a letter apologizing for its error and agreeing with the USPTO's calculation[28]. Thus, at least the government is coming up to speed on computing deadlines.

The Prilosec OTC PTE applicant petitioned the USPTO director, alleging that the USPTO's "new" method of calculating the 60-day deadline would cause "unduly prejudicial and detrimental consequences for the Applicant"[29]. In its petition, the applicant highlighted the extensive confusion surrounding the 60-day deadline and alleged that 13 patents have been granted extensions despite applications being filed after the 60-day deadline, as calculated by the USPTO's method. Although several PTE applications were indeed filed late and then improperly granted, this is a tangled situation that may only spread the grief around rather than resolve it for the Prilosec OTC PTE applicant, especially because the applicant's petition does not address the clear language in the statute stating that the 60-day period begins on the date of regulatory approval. In

Table 2 Drugs that have forfeited their PTE eligibility by missing the filing deadline

Product name	Applicant	Date of FDA approval	PTE appl. filing date	True deadline	Mistaken party	PTE status
Angiomax	NDA holder	Dec. 15, 2000	Feb. 14, 2001	Feb. 12, 2001	Applicant and USPTO (initially)	Denied
A180	Patentee	Sept. 20, 2002	Nov. 19, 2002	Nov. 18, 2002	Applicant, USPTO and FDA (initially)	Denied
Prilosec OTC	Patentee	June 20, 2003	Aug. 19, 2003	Aug. 18, 2003	Applicant and FDA (initially)	Denied
Symbicort	Patentee	July 21, 2006	Sept. 19, 2006	Sept. 18, 2006	Applicant	Denied, reconsideration pending

response to this petition, the USPTO issued a "final agency decision" denying PTE based on the same grounds[30].

In the Symbicort case[31], the FDA got it right on the first try and identified the application as untimely[32]. The USPTO later denied PTE in June 2008 for this and other reasons[33]. In its decision, the USPTO expressed surprise at the applicant's miscalculation, stating, "It is unclear how Applicant, who specifically correctly indicated that the first day of the sixty-day period 'began on July 21, 2006,' calculated that the end point of the sixty-day period was any day other than September 18, 2006"[34]. Hence, the USPTO may still be unaware of the extent of these errors. In its request for reconsideration, the applicant provided essentially the same arguments and supporting information presented in the Prilosec OTC case, but referenced a 14th patent for which a PTE was granted even though the application was filed after 60 days[35].

Consequences of improperly granted PTE

As of this writing, at least 13 patents have been improperly extended based on late-filed PTE applications, four of which have not yet expired. Because some PTE applications are not viewable through the PAIR database, it is possible that other granted applications and/or pending applications were untimely, but the USPTO and/or FDA may have failed to spot the errors. A range of issues could arise from such errors, apart from the patent being limited to its original expiration date and thereafter unable to protect the patented product. An improperly extended patent listed in the Orange Book could also give rise to liability by unjustly deterring a generic drug company from pursuing a competing product[36], resulting in higher drug prices for a longer period of time. Other issues include possible license agreement violations, where royalties are normally payable for the life of a licensed patent, including any extensions, or patent misuse if a patent owner seeks to enforce an expired patent or knowingly continues to collect royalties past the expiration date.

Other PTE pitfalls

Another avoidable pitfall arises in calculating the duration of a requested extension. The extension period is based on the regulatory review period—that is, the testing and approval phases preceding FDA approval. The testing phase starts on the effective date of the Investigational New Drug (IND) application (which is required to conduct drug testing in humans) and ends on the filing date of the New Drug Application (NDA). The approval phase begins on the NDA filing date and ends with the grant of FDA approval. In general, a PTE corresponds to half the number of days in the testing phase plus the total number of days in the approval phase, excluding time periods when the applicant failed to act with "due diligence"[37]. Although a lack of diligence during drug development can shorten the extension, lawyers should be mindful not to shortchange a PTE applicant unnecessarily.

Yet another tricky point comes in determining the effective start date of the testing period. Drug companies often have several drugs in their pipelines and tend to move resources from one to another, particularly at early stages when it is unclear whether a target candidate will be of sufficient commercial value to proceed. Drug development can also move from one company to another as dictated by business interests, causing the work to proceed in fits and starts. All of these activities and their time periods should be assessed to determine the longest possible extension for which the patent qualifies.

A further PTE nuance relates to the NDA filing date for the drug that is the subject of a PTE application. This date is important because it marks both the end of the testing phase and the start of the approval phase. Forced by the statutory language, the FDA counts this date twice in making a PTE determination—once as the end date of the testing phase, and again as the start date of the approval phase[38]. For drugs that generate millions of dollars per day, excluding this date in a request for PTE, or suffering from its exclusion if the FDA makes such an error, could be very costly.

PTE practice points

Although it appears that the USPTO and FDA are both becoming more vigilant about the calculation, many PTE applicants and their lawyers are still making mistakes. Unless and until the law is changed, the USPTO, FDA and especially lawyers filing PTE applications should be extremely careful when calculating the 60-day deadline. Clearly, lawyers should file these applications well before the deadline and set their target filing date at least five days before the presumed deadline to avoid the severe consequences of miscalculation. Generic companies on the other hand should always check whether the PTE of a listed patent was properly granted, as the resulting ineligibility could provide them at a minimum with an easy defense (an expired patent cannot be infringed) or even with some ammunition to fight against the brands (e.g., patent misuse).

In assessing due diligence and calculating the duration of a possible PTE, it is particularly important for the PTE application to identify all significant drug development activities and milestones, regardless of any duration of inactivity that occurs between activity dates. For instance, if activity started but was then stalled for several years, the initial start date should be disclosed to the FDA, as well as any noteworthy activities that may have occurred during the period of relative inactivity. It may also be possible for nontechnical business development practices to be considered evidence of diligent efforts, even if laboratory or clinical testing was not occurring at the time. The applicant is not required to prove due diligence in the first instance, probably because the FDA can rely on its own corresponding records, but providing a complete factual disclosure of all the relevant activities is in the applicant's best interest as the FDA may calculate a longer regulatory review period, corresponding to a longer PTE, which will be granted if unchallenged for 180 days after publication of the FDA's calculation (or if the PTE applicant prevails against any such challenge). Notably, the granted length of extension will also be free from later challenge in court, foreclosing a claim of inadequate due diligence as a possible noninfringement defense[39]. Providing a full description of the relevant facts also complies with the duty of disclosure, which requires the applicant to disclose all information that would be considered "important in determinations to be made in the patent term extension proceeding," including all "material information adverse to a determination of entitlement to the extension sought"[40]. Such full disclosure can support a good faith request for a longer extension period and might even deter a third party from challenging the extent of PTE granted[41].

For determining the effective start date of the testing period, the PTE applicant should look to the earliest activities related to the product at issue to determine the earliest possible start date. For instance, the testing phase can start when the IND becomes effective (typically 30 days after the IND is filed). If an earlier IND was filed to test a different indication, further investigation should be made to determine if any of the earlier activity (e.g., safety testing) would permit reliance on the earlier start date in calculating the extension period.

For generic drug companies, it is critical to confirm that any grant of PTE was properly made (and was based on a timely filed PTE application) before certifying against an extended patent listed in the Orange Book. If the PTE application was untimely and the original expiration has not yet passed, the generic company might certify that the generic drug will not be marketed before the original expiration date (a "Paragraph III Certification"), which may not require a very long wait. If the original patent term has lapsed, the generic

might be able to certify that the listed patent has expired (a "Paragraph II Certification"), but it is more likely that the FDA would require the generic in this instance to certify that the patent would not be infringed on the grounds that PTE was improperly granted[42] or is unenforceable because an expired patent cannot be enforced (a "Paragraph IV Certification"). This last type of certification can be extremely valuable because the first abbreviated NDA applicant seeking to gain approval of a generic drug can be rewarded with 180 days of market exclusivity if it includes a Paragraph IV Certification in its application.

COMPETING INTERESTS STATEMENT
The author declares competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturebiotechnology/.

1. Johnson & Johnson Fourth Quarter and Year End 2008 Earnings Meeting (Jan. 20, 2009) <http://files.shareholder.com/downloads/JNJ/531545955x0x264950/c2d48b4a-d107-48ee-b793-0f17ab5033ae/4Q08_WCWeldon_V10_Web.pdf>
2. Pfizer Reports Fourth-Quarter And Full-Year 2008 Results And 2009 Financial Guidance (Jan. 26, 2009) <http://media.pfizer.com/files/investors/presentations/q4performance_january012609.pdf>
3. AstraZeneca PLC Fourth Quarter and Full Year Results 2008 (Jan. 29, 2009) <http://www.astrazeneca.com/_mshost3690701/content/resources/media/investors/AZN-Q4-2008/q4-results-2008-figures.pdf>
4. 35 USC §271.
5. 35 USC §156; 21 CFR §60; Drug Price Competition and Patent Term Restoration Act of 1984, tit. II, Pub. L. No. 98–417, 98 Stat. 1585, 1598 (1984) (as amended Dec. 8, 2003).
6. 21 USC 355(b).
7. Electronic Orange Book (FDA, Washington, DC;). <http://www.fda.gov/cder/ob>
8. 21 USC §355(j); 21 USC §355(b)(2)
9. 35 USC §156(d)(1).
10. Online date calculators follow this convention as well. See, e.g. Date Calculator <http://www.timeanddate.com/date/dateadd.html>
11. USPTO Patent Application Information Retrieval (PAIR) database. <http://portal.uspto.gov/external/portal/pair>
12. A PTE applicant must identify the filing deadline in its application. 37 CFR §1.740(a)(5).
13. Manual of Patent Examining Procedure §2753 (July 2008 revision).
14. *See* Application Pursuant to 35 USC §156(d)(1) and 37 CFR §1.740 For Extension of Patent Term for US Patent No. 5,196,404. <http://portal.uspto.gov/external/portal/pair>. This PTE applicant apparently read the statute as reciting a two-month deadline based on its identificaiton of February 15, 2001 as the sixtieth day.
15. Letter from USPTO to FDA for US Patent No. 5,196,404 (Mar. 2, 2001) Notice of Final Determination (Nov. 20, 2001). <http://portal.uspto.gov/external/portal/pair>
16. Decision Denying Application for Patent Term Extension for US Patent No. 5,196,404 at p. 7, n. 3 (Apr. 26, 2007). <http://portal.uspto.gov/external/portal/pair>
17. 888 F.2d 826 (Fed. Cir. 1989).
18. HR 5120, 109th Cong. (2006).
19. HR 6344, 110th Cong. §4(a)(1) (as passed by House of Representatives, June 23, 2008).
20. Id. at §4(b)(2)(B)
21. Cong. Rec. H5811–H5816 (daily ed. Jun. 23, 2008) (statement of Rep. Conyers).
22. *See* Application for Extension of the Term of US Patent No. 4,861,779 Under 35 USC §156 (Nov.19, 2002). <http://portal.uspto.gov/external/portal/pair>
23. Letter from FDA to USPTO [re: US Patent No. 4,861,779] (Jul. 16, 2003). <http://portal.uspto.gov/external/portal/pair>
24. Order to Show Cause [re: US Patent No. 4,861,779] (April 4, 2008). <http://portal.uspto.gov/external/portal/pair>
25. *See* Application for Extension of Patent Term Under 35 USC §156 [re: U.S. Patent No. 5,817,338] (Aug. 19, 2003). <http://portal.uspto.gov/external/portal/pair>. This PTE application was also rejected on other grounds.
26. Letter from FDA to USPTO [re: US Patent No. 5,817,338] (Oct. 19, 2004). <http://portal.uspto.gov/external/portal/pair>
27. Letter from USPTO to FDA [re: US Patent No. 5,817,338] (Apr. 1, 2008). <http://portal.uspto.gov/external/portal/pair>
28. Letter from FDA to USPTO [re: US Patent No. 5,817,338] (Oct. 21, 2008). <http://portal.uspto.gov/external/portal/pair>
29. Petition to the Director (37 CFR §1.181) [re: US Patent No. 5,817,338] (May 30, 2008). <http://portal.uspto.gov/external/portal/pair>
30. Denial of Patent Term Extension Application [re: US Patent No. 5,817,338] (Dec. 16, 2008). <http://portal.uspto.gov/external/portal/pair>
31. *See* Application for Extension of Patent Term Under 35 USC §156 [re: U.S. Patent No. 5,674,860] (Sept. 19, 2003). <http://portal.uspto.gov/external/portal/pair>. This PTE application was also rejected on other grounds.
32. Letter from FDA to USPTO [re: US Patent No. 5,674,860] (Dec. 6, 2007). <http://portal.uspto.gov/external/portal/pair>
33. Notice of Final Determination – Ineligible [re: US Patent No. 5,674,860] (Jun. 13, 2008). <http://portal.uspto.gov/external/portal/pair>
34. Id. at p. 5.
35. Request for Reconsideration of Final Determination [re: U.S. Patent No. 5,674,860] (Dec. 16, 2008). <http://portal.uspto.gov/external/portal/pair>
36. Cf. In re Buspirone Antitrust Litigation, 185 F. Supp. 2d. 340 (improperly listing a patent in the Orange Book associated with potential antitrust liability).
37. 35 USC §156(c)(1).
38. 37 CFR §1.775(c).
39. 35 USC §282.
40. 37 CFR §1.765(a).
41. MPEP §2753; 35 USC §156(d)(2)(B).
42. 35 USC §282.

Recent patent applications in genomic assays

Patent number	Description	Assignee	Inventor	Priority application date	Publication date
WO 2008003100, EP 2038074	A method for functionalizing a surface with sites to form, e.g., a peptide array for protein detection, involving attaching two sets of molecules of predetermined concentrations to sites and treating density-variation surface sites with reactive molecules.	Gao X, Hong A, Yu P, Zhang X, Zhou X, Zhu Q	Gao X, Hong A, Yu P, Zhang X, Zhou X, Zhu Q	6/29/2006	1/3/2008, 3/25/2009
WO 2009021141	A method for designing a microarray for DNA methylation analysis comprises generating probe sequences corresponding to the CpG islands and generating discrete genomic regions.	Johns Hopkins University (Baltimore, MD, USA)	Feinberg A, Irizarry R	8/7/2007	2/12/2009
WO 2008034622	A method of detecting and/or quantifying expression of a target protein that permits target identification without the need for complex biochemical assays, thus serving as a genome-wide approach for target identification.	Pasteur Institute Korea (Seoul)	Emans N, Nehrbass U	9/20/2006	8/21/2008
US 20080064040	A new assay panel comprising fluorescence or luminescence assays; useful for the detection of gene or protein silencing, gene or protein inactivation, or gene or protein activation in an intact cell.	Odyssey Thera (San Ramon, CA, USA)	Belisle B, Lamerdin JE, MacDonald ML, Watson MSW, Westwick JK	5/30/2003	3/13/2008
WO 2007070560, EP 1977005	A method of identifying an imprinted gene comprising analyzing the chromatin structure of the subject's genome to establish a map of the subject's chromatin structure.	Bjornsson HT, Nimblegen Systems (Madison, WI, USA), Feinberg AP, Green RD	Bjornsson HT, Feinberg AP, Green RD	12/13/2005	8/23/2007
WO 2005020886, US 20060111848	A computational method for identifying proteins in a pathogen suitable as a target to detect a therapeutic agent, by determining a computationally phenotypic change in the pathogen by loss of proteins and identifying the ranking order of proteins.	New England Biolabs (Ipswich, MA, USA)	Carlow CKS, Foster J, Kumar S, Zhang Y	6/27/2003	3/10/2005, 5/25/2006
US 20040241636, WO 2005001115	Annotating gene or protein function by measuring, for example, the quantity of a test protein in the first and second population of cells contacted with a test or control reagent, comparing results and identifying the proteins linked to the test gene or protein.	Belisle B, Lamerdin JE, MacDonald ML, Michnick SWW, Odyssey Thera (San Ramon, CA, USA), Westwick JK	Belisle B, Lamerdin JE, MacDonald ML, Michnick SWW, Westwick JK	5/30/2003	12/2/2004, 1/6/2005
WO 2003064701, EP 1470254	Identifying a reliable diagnostic, prognostic or staging marker for phenotypic conditions characterized by altered DNA methylation, for example, cancer, comprising obtaining a set of at least two biological samples in each case having genomic DNA.	Epigenomics (Berlin)	Schweikhardt RG, Sledziewski AZ	1/30/2002	8/7/2003, 10/27/2004
US 20040115722	A method of producing a biopolymeric array useful for diagnostic or screening procedures or in gene expression analysis, comprising immobilizing different probes for respective targets on a solid support.	Caren MP, Corson JF, Kronick M, Leproust EM, Peck BJ	Caren MP, Corson JF, Kronick M, Leproust EM, Peck BJ	10/25/2002	6/17/2004
WO 2004015080	A method of detecting alternative splicing of a first target sequence comprising hybridizing probes to target sequences without prior amplification to form hybridization complexes; useful in diagnostic medicine.	Illumina (San Diego, CA, USA), Regents of the University of California (Oakland, CA, USA)	Fan J, Fu X	8/9/2002	2/19/2004

Source: Thomson Scientific Search Service. The status of each application is slightly different from country to country. For further details, contact Thomson Scientific, 1800 Diagonal Road, Suite 250, Alexandria, Virginia 22314, USA. Tel: 1 (800) 337-9368 (http://www.thomson.com/scientific).

Shining light on a new class of hydrogels

Steven M Jay & W Mark Saltzman

Addition of a photodegradable group to the backbone of synthetic hydrogels enables real-time control of the material's chemical and physical properties.

Hydrogels are essential ingredients in current strategies for tissue engineering and cell delivery owing to their physical properties, which mimic aspects of natural extracellular matrix. As described recently in *Science*, Kloxin *et al.*[1] have devised a hydrogel-forming polymer whose physicochemical properties can be controlled by light. This report heralds the arrival of a new class of hydrogels with physical and chemical properties that can be dynamically controlled.

Hydrogels were used in the earliest reports on tissue engineering nearly 30 years ago. Initially formed from proteins derived from natural extracellular matrix, they served to entrap cells and facilitate tissue formation[2]. Hydrogels of natural polymers, such as collagen and fibronectin, remain important materials in tissue engineering[3], but investigators are relying increasingly on hydrogels made of artificial materials to take advantage of the versatility and functionality provided by synthetic polymer chemistry[4]. Over the past few decades, scientists and engineers have created materials to enhance hydrogel properties, moving them from their beginnings as inert, biocompatible scaffolds to gels that respond to electrical and mechanical stimuli, can be assembled entirely from molecularly defined components with multiple functionalities and react to cellular microenvironments by degrading or releasing drugs or biological moieties[5].

Despite these advances, however, active, time-dependent, high-resolution control of hydrogel properties has been lacking. Kloxin *et al.*[1] have addressed this deficiency by developing new materials that allow the physical and chemical properties of cell-bearing hydrogels to be modified after gelation without cytotoxicity. Previous work[6] established the utility of hydrogels that contain photolabile groups attached to a photostable backbone for modulating three-dimensional, directed tissue growth. Kloxin *et al.*[1] extended this concept with an important chemical advance. By incorporating a photodegradable group into the polymeric backbone, they rendered the gel network photocleavable, enabling gross modulation of gel characteristics even in the presence of entrapped cells (**Fig. 1a**). When these new materials were combined with photolithography, it was possible to manipulate gel features rapidly in three dimensions with micrometer resolution (**Fig. 1b**).

This dynamic control of both physical and chemical properties—coupled with the ability to release pendant functional groups or molecules—has significant implications for tissue engineering. Although the ultimate goal of this field is replacement of failing organs or tissues *in vivo*, re-creation of complex tissue structure *in vitro* is useful for discovering how cell-matrix interactions dictate differentiation and tissue development[7]. Hydrogels provide a setting to evaluate such interactions in three dimensions, but the lack of facile, real-time control of the chemistry of cell microenvironments has been a limitation in mimicking the dynamic interaction between cells and extracellular matrix that occurs *in vivo*[8]. Manipulation of bulk hydrogel properties typically requires addition of chemical or biological entities. The techniques reported by Kloxin *et al.*[1] allow for observation of cell behavior in response to real-time changes in the local physical properties of the gel, such as stiffness, simply by adding light.

For example, the authors used photodegradation to decrease the density of cross-linking within gels that entrapped human mesenchymal stem cells, facilitating spreading of cells and altered function (**Fig. 1a**). They also showed that it is possible to induce directed migration of embedded cells into micron-scale three-dimensional structures arbitrarily defined by photodegrading the gel with a laser scanning microscope (**Fig. 1b**). Although not explicitly demonstrated yet, it is easy to imagine the application of this technique to directed migration and spreading of other cell types, such as neurons[6], to create tissue structures.

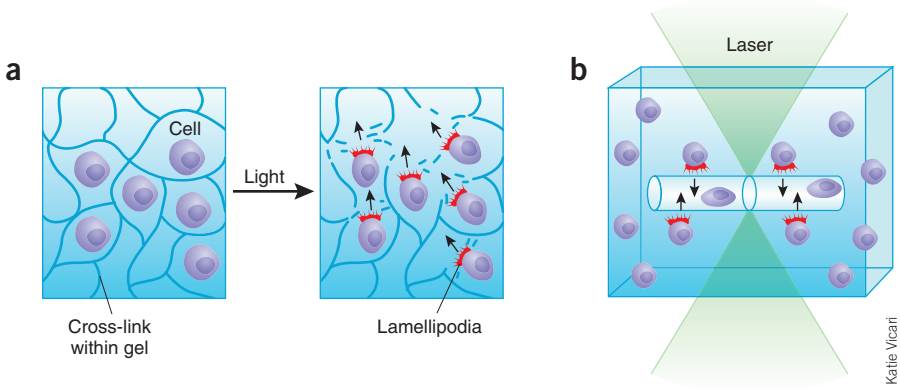


Figure 1 Photodegradable hydrogels allow for the dynamic manipulation of the environment of cells. (**a**) Local gel degradation enables spreading of entrapped cells (arrows indicate direction of spreading). (**b**) Controlled photodegradation of micron-sized channels facilitates migration of encapsulated cells.

Steven M. Jay and W. Mark Saltzman are in the Department of Biomedical Engineering, Interdepartmental Program in Vascular Biology and Therapeutics, Yale University, New Haven, Connecticut, USA.
e-mail: mark.saltzman@yale.edu

Furthermore, using methods for attachment and photocleavage of pendant peptide sequences, Kloxin *et al.*[1] show that photoinduced removal of the adhesion sequence Arg-Gly-Asp-Ser from a cross-linked gel microenvironment after 10 days of human mesenchymal stem cell culture results in enhanced chondrogenic differentiation, as measured by an increase in glycosaminoglycan production associated with decreased CD105 expression. This manipulation mimics the reported downregulation of the adhesion protein fibronectin by chondrocyte-destined human mesenchymal stem cells[9]. This experiment serves as a bellwether for future studies, which likely will examine the effects of dynamic changes in ligand presentation on cell adhesion, migration and differentiation

Dynamic control of tissue regeneration *in vivo* remains a critical challenge in tissue engineering. It is not clear that the on-demand alteration of gel properties reported by Kloxin *et al.*[1] will be directly translatable to *in vivo* applications, but the use of light as the source of control is promising. Many common medical procedures involve exposure to light at various wavelengths, and laser and optical imaging technologies are constantly improving. In this new report, the authors make clear that exposure to irradiation sufficient to substantially photodegrade hydrogels is not harmful to entrapped human mesenchymal stem cells .

The techniques allow for observation of cell behavior in response to real-time changes in the local physical properties of the gel, such as stiffness, simply by adding light.

What if these techniques could be used *in vivo*? Exciting possibilities for this technology are easy to imagine. For example, drugs could be repeatedly and controllably released in response to a tissue-penetrating light beam from subcutaneously implanted gel depots. A similar concept could be applied to cell delivery. The sustained deployment of cells from a biomaterial scaffold based on the formation of growth-factor diffusional gradients has recently been reported to be a more effective means of cell therapy than conventional scaffold-based cell administration or bolus cell injection[10]. The method described by Kloxin *et al.*[1] could be used in a similar mode, allowing physicians to noninvasively dictate when and how many cells are released from an implanted reservoir. Such a therapy would represent an unprecedented opportunity to incorporate feedback control into cell and drug delivery, increasing therapeutic potential while limiting risk to the patient.

The eventual utility of dynamically controllable, photodegradable hydrogels will depend on factors that are currently difficult to predict, including cost, versatility and *in vivo* safety. However, considering the long-running clinical application of numerous modes of photomedicine, such as photodynamic therapy[11], it is reasonable to believe that the technology put forth by Kloxin *et al.*[1] will one day find clinical application.

1. Kloxin, A.M., Kasko, A.M., Salinas, C.N. & Anseth, K.S. *Science* **324**, 59–63 (2009).
2. Bell, E., Ehrlich, H.P., Buttle, D.J. & Nakatsuji, T. *Science* **211**, 1052–1054 (1981).
3. Jay, S.M., Shepherd, B.R., Bertram, J.P., Pober, J.S. & Saltzman, W.M. *FASEB J.* **22**, 2949–2956 (2008).
4. Lutolf, M.P. & Hubbell, J.A. *Nat. Biotechnol.* **23**, 47–55 (2005).
5. Kopecek, J. *Biomaterials* **28**, 5185–5192 (2007).
6. Luo, Y. & Shoichet, M.S. *Nat. Mater.* **3**, 249–253 (2004).
7. Griffith, L.G. & Swartz, M.A. *Nat. Rev. Mol. Cell Biol.* **7**, 211–224 (2006).
8. Yamada, K.M. & Geiger, B. *Curr. Opin. Cell Biol.* **9**, 76–85 (1997).
9. Tavella, S. *et al. J. Cell Sci.* **110**, 2261–2270 (1997).
10. Silva, E.A., Kim, E.S., Kong, H.J. & Mooney, D.J. *Proc. Natl. Acad. Sci. USA* **105**, 14347–14352 (2008).
11. Dolmans, D.E., Fukumura, D. & Jain, R.K. *Nat. Rev. Cancer* **3**, 380–387 (2003).

Cause and express

Leonid Kruglyak & John D Storey

Biological validation of a cadre of new obesity genes supports the power of studies that exploit 'expression quantitative trait loci'.

The biological basis of complex traits such as obesity has long been sought among the varying genotypes of individuals. In a new twist on classical genetic mapping, geneticists have begun to use microarrays to link genomic regions to mRNA abundance[1] and then to candidate genes responsible for complex traits or diseases[2]. Reporting in *Nature Genetics*, Drake and colleagues[3] have carried out validation studies of genes previously implicated in obesity through analysis of genomic loci that influence gene expression levels—often called expression quantitative trait loci, or eQTLs. Their results support the utility of these new genetic methods.

In eQTL studies, two types of information are collected from genetically diverse individuals, such as progeny from a cross: each individual is genotyped with molecular markers, and gene expression is measured in cells or tissues of interest. These data are jointly analyzed to identify regions of the genome that contain polymorphisms affecting gene expression. The expression level of each gene is treated as a 'quantitative trait'—that is, a phenotypic measurement that shows a continuous range of variation among individuals.

Classical genetic mapping typically considers quantitative traits that are organismal phenotypes, such as blood pressure or body weight, and maps them to variation-containing regions of the genome, called quantitative trait loci. eQTL studies extend this approach in two key respects: microarrays allow thousands of expression quantitative traits to be studied simultaneously, and expression traits serve as an intermediate between DNA sequence variation and organismal phenotypes[4]. Such studies have provided insights into the genetic basis of phenotypic variation[5] and the structure of regulatory networks[6,7].

To predict disease-causing genes, researchers have begun to combine gene expression and genotype data with clinically relevant phenotypes such as disease status, physical traits and biomarkers[2,8]. Although the strategy of using genetics to infer causality among multiple traits has been in the literature for several decades[9,10], only recently has the inclusion of 'gene expression traits' into this framework been possible.

The first step in this process is to identify a gene expression trait and a clinical trait (such as abdominal obesity) that are genetically linked to a common locus. There are several relationships possible between the expression trait and clinical trait based on this common linkage (**Fig. 1a**). For example, the locus may influence the two traits independently (by means of either the same or different polymor-

Leonid Kruglyak is at the Lewis-Sigler Institute for Integrative Genomics, Department of Ecology and Evolutionary Biology and the Howard Hughes Medical Institute and John D. Storey is at the Lewis-Sigler Institute for Integrative Genomics and the Department of Molecular Biology, Princeton University, Princeton, New Jersey, USA.
e-mail: leonid@genomics.princeton.edu or jstorey@princeton.edu

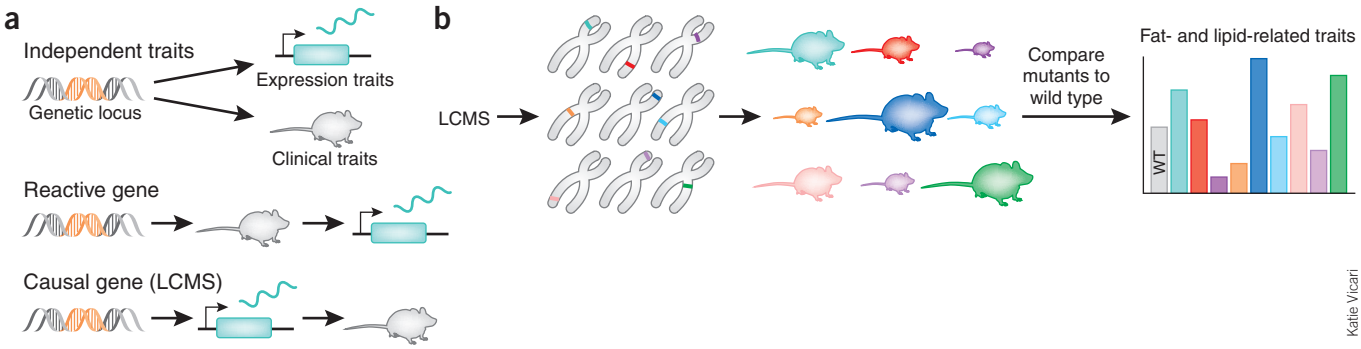


Figure 1 Expression quantitative trait loci (eQTLs) are useful in understanding the molecular basis of complex diseases. (**a**) A genetic locus may independently affect both a gene expression trait and a clinical trait. Alternatively, gene expression changes may result from genetically influenced phenotypes. LCMS is a method for identifying causal relationships—that is, genetic polymorphisms that affect gene expression, which in turn affect a clinical trait. (**b**) Drake and colleagues[3] constructed or obtained genetically modified mouse strains for nine genes identified as causal for obesity by LCMS. Measurements of obesity in these strains and in wild-type controls validated the causal role of eight gene.

phisms), in which case neither trait affects the other. Alternatively, the expression trait may be 'reactive'—that is, the locus may influence the clinical trait, which in turn influences the gene expression trait. Finally, the locus may influence the gene expression trait, which in turn influences the clinical trait. In these cases, the gene is predicted to be 'causal' for the clinical trait.

In a previous study[11] that relied on this last approach, ~100 genes had been predicted to be causal for abdominal obesity in the mouse. The method used is called likelihood-based causal model selection (LCMS). LCMS does not apply a statistical significance test for causality. Rather, it considers many different models and chooses the one that best fits the data. A challenge in applying LCMS is that it can predict thousands of genes to be causal, most of which are probably false positives. Another challenge is that unknown or unmeasured factors may influence both gene expression and the clinical trait, giving the false appearance that a causal relationship exists.

To narrow the list of genes deemed causal by LCMS to a manageable size for experimental validation and to reduce the risk of false positives, Drake and colleagues[3,11] applied additional criteria to select a candidate list of nine genes. An alternative approach would have been to take advantage of the partial 'mendelian randomization' of both the expression trait and the clinical trait, which occurs because of the random segregation of alleles at the locus to which they are both genetically linked. If this approach had been used, it would have been possible to perform a direct significance test for causality, thereby allowing one to guard against spurious findings due to confounding hidden factors and to avoid the need for additional manual gene filtering[6].

To validate the nine genes predicted to be causal for obesity, Drake and colleagues[3] constructed or obtained genetically modified mouse strains in which one of the nine genes was either knocked out or overexpressed (**Fig. 1b**). They compared a large number of obesity-related traits in male and female mice from these strains to wild-type controls. Eight of the nine genes showed some effect on one or more of the traits. Interestingly, for two of the genes, the direction of the effect was different between males and females. Drake and colleagues[3] also compared genome-wide gene expression between each genetically modified mouse strain and its corresponding control. The effects on gene expression were modest, with few individual genes showing statistically significant differences. However, the authors observed some significant enrichment of obesity-related metabolic pathways and partial overlap of the overall expression signatures.

Elucidating the networks and pathways relevant to disease is a first step to identifying specific genes whose products can serve as drug targets.

Drake and colleagues[3] assumed that genes correlated with the trait can be either causal for the trait or reactive to the trait, but not both. In the future, it would be interesting to perform similar assays on a sizable group of negative control genes—specifically, those that are correlated with the clinical trait but are not found to be causal by LCMS. Such a study may reveal that the situation is not so black and white. That is, genes may not have mutually exclusive causal or reactive relationships to a trait but instead may be linked by means of an intricate network with feedback loops, in which perturbation of any component leads to small changes in clinical traits and subtle alterations of global gene expression patterns.

The ultimate goal of combining eQTL and clinical studies in humans and mammalian disease models is to develop new treatments. Elucidating the relevant networks and pathways is a first step to identifying specific genes whose products can serve as drug targets. Both computational predictions and validation studies such as those performed by Drake and colleagues[3] are useful for sorting out which of the many gene products are likely to have a large desired effect when inhibited. From this perspective, the observation that some genes showed opposite effects in males and females raises the possibility of sex-specific treatments. We await the next steps in which a therapeutic is used to target gene products identified via an eQTL approach and achieve a desired outcome, first in an animal model and then in patients.

1. Brem, R.B., Yvert, G., Clinton, R. & Kruglyak, L. *Science* **296**, 752–755 (2002).
2. Cookson, W., Liang, L., Abecasis, G., Moffatt, M. & Lathrop, M. *Nat. Rev. Genet.* **10**, 184–194 (2009).
3. Yang, X. *et al. Nat. Genet.* **41**, 415–423 (2009).
4. Rockman, M.V. & Kruglyak, L. *Nat. Rev. Genet.* **7**, 862–872 (2006).
5. Brem, R.B. & Kruglyak, L. *Proc. Natl. Acad. Sci. USA* **102**, 1572–1577 (2005).
6. Chen, L.S., Emmert-Streib, F. & Storey, J.D. *Genome Biol.* **8**, R219 (2007).
7. Zhu, J. *et al. PLOS Comput. Biol.* **3**, e69 (2007).
8. Yvert, G. *et al. Nat. Genet.* **35**, 57–64 (2003).
9. Katan, M.B. *Lancet* **1**, 507–508 (1986).
10. Davey Smith, G. & Ebrahim, S. *Int. J. Epidemiol.* **32**, 1–22 (2003).
11. Schadt, E.E. *et al. Nat. Genet.* **37**, 710–717 (2005).

Structure of a multidrug transporter

Michael M Gottesman, Suresh V Ambudkar & Di Xia

Crystal structures of a mammalian multidrug efflux pump bound to peptide inhibitors may reveal drug-binding sites.

Multidrug transporters are the bane of pharmacologists as they diminish the efficacy of many drugs by pumping them out of bacteria and mammalian cells[1]. The multidrug efflux pump P-glycoprotein (P-gp), for example, contributes to multidrug resistance in about half of human cancers. In a recent paper in *Science*, Aller *et al.*[2] reported X-ray crystal structures of mouse P-gp to 3.8 Å resolution—the highest-resolution structures of a mammalian multidrug transporter to date. The three structures show P-gp alone and P-gp bound to two stereoisomers of a novel cyclic hexapeptide inhibitor, all in the absence of ATP. These results reinforce many of the conclusions derived previously from detailed biochemical studies and illustrate the pitfalls of using static structures to understand conformationally dynamic molecules such as ATP-binding cassette (ABC) transporters.

The mammalian P-gp transporters (encoded by the *MDR1* (*ABCB1*) gene in the human and the *mdr1a* (*abcb1a*) and *mdr1b* (*abcb1b*) genes in the mouse) were first cloned in 1986 (refs. 3,4), and their essential physiological, biochemical and pharmacologic features have been well described[5]. However, controversy remains over how P-gp recognizes hundreds of different hydrophobic drugs and pumps them out of the cell and how similar it is to the 47 other known members of the human ABC transporter family. High-resolution crystal structures of P-gp would be invaluable for investigating these questions and for designing therapeutic P-gp antagonists.

Although the structures solved by Aller *et al.*[2] lack the high resolution needed to settle most of the remaining controversies, they do provide useful new information. The essential features are a large, 6,000 Å^3 internal cavity to which hydrophobic drugs can bind in various orientations and at different locations, as well as two portals in the inner leaflet part of the protein that allow entry of hydrophobic drugs from the lipid bilayer. The structures support the "hydrophobic vacuum cleaner" model first proposed to explain the observation that P-gp can remove hydrophobic drugs directly from the membrane[6]. The drug-binding cavity is composed of parts of the transmembrane segments from both halves of P-gp, especially transmembrane helices 4, 5 and 6 and 10, 11 and 12, as predicted from previous photoaffinity labeling studies[7,8] and cross-linking analyses[9]. Although not of sufficiently high resolution to discern individual side chains, the structures make many clear predictions about residues likely to be involved in drug binding that can and will be tested by mutational analysis. Interestingly, P-gp crystallizes into identical crystal forms in the absence of any inhibitor (PDB:3G5U) and in the presence of one (PDB:3G60) or two inhibitors (PDB:3G61) in the central cavity. The three structures display few changes in crystal cell parameters, the overall conformation of the protein (0.6 Å root-mean-square deviations for all Cα atoms) and the local environment of the binding sites, except for a few changes in residues that directly contact the bound inhibitor.

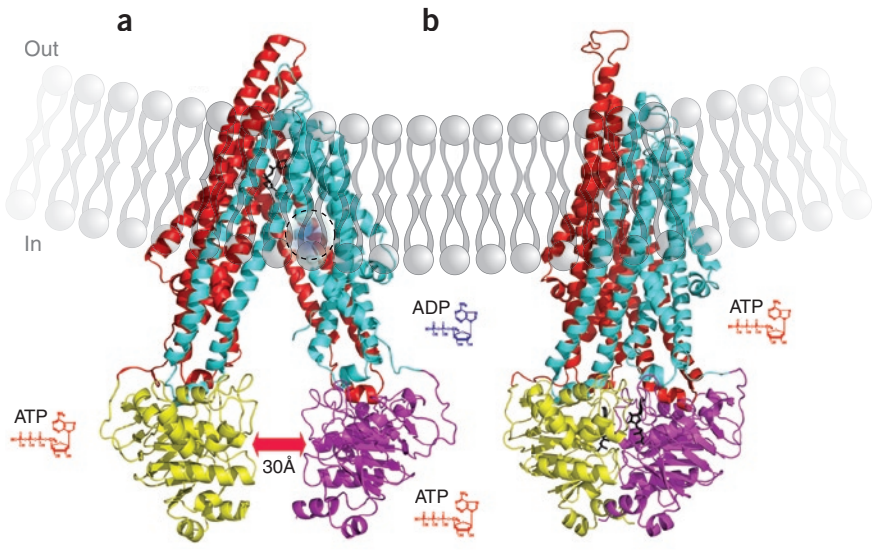


Figure 1 Ribbon renditions of the structure of P-glycoprotein embedded in a membrane bilayer. ATP (red) and ADP (blue) molecules are shown present in cytoplasm. (**a**) Mouse P-gp at 3.8 Å resolution (PDB: 3G60) as described by Aller *et al.*[2]. The N-terminal transmembrane domain and nucleotide-binding domain are shown in red and yellow, respectively, and the C-terminal transmembrane domain and nucleotide-binding domain are in cyan and magenta, respectively. The inhibitor QZ59-SSS bound in the transmembrane region is shown as a stick model in black. The red double-arrow indicates a distance of 30 Å between the two 'nucleotide-free' nucleotide-binding domains. The dashed black circle indicates an area proposed to be one of the portals for entry of substrates or modulators directly from the membrane. (**b**) Model of human P-gp based on the structure of *S. aureus* SAV1866 bound to ADP and open to the extramembrane space (PDB: 2HYD) from ref. 10. Two bound ADP molecules are shown in black. Color code follows that in **a**.

Several features of the model will undoubtedly stimulate more questions than they answer. To crystallize unglycosylated mouse P-gp purified from the yeast *Pichia pastoris*—an accomplishment that had eluded all others in the field—Aller *et al.*[2] had to eliminate ATP, ATP analogs and magnesium from the crystallization medium. The structure (**Fig. 1a**) shows the two halves of the potential ATP-binding sites 30 Å apart—a considerable distance given that these regions must interact to bind ATP during the catalytic cycle.

The authors speculate that the structure in **Figure 1a** represents mouse P-gp in a "pretransport state" and that binding of ATP during the catalytic cycle probably results in dimerization of the ATP sites. However, because ATP concentrations in the cell (3–5 mM) always far exceed the affinity of the transporter for ATP (K_{mATP} 0.3–1 mM), it seems unlikely that P-gp

Michael M. Gottesman, Suresh V. Ambudkar and Di Xia are at the Laboratory of Cell Biology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA.
e-mail: mgottesman@nih.gov

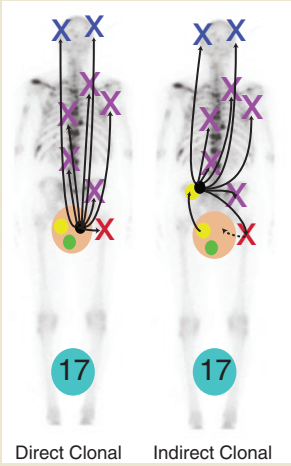
ever exists in the cell in a 'nucleotide-free' state. Thus, the structure may represent a crystallization artifact or a nonfunctional conformation that has only very transient existence. The lack of significant conformational changes in the transmembrane domains upon inhibitor binding also supports this concern.

The structure of Aller *et al.*[2] differs in substantial ways from previous models of P-gp (**Fig. 1b**) based on the so-called 'closed conformation' of an ABC transporter from *Staphylococcus aureus* SAV1866 (ref. 10). In addition to the above-mentioned 30 Å distance between the two nucleotide-binding domains and the lack of nucleotide binding, the other major difference is that the Aller *et al.*[2] structure has a transmembrane domain conformation that is open to the inside, whereas the model based on SAV1866 has a conformation open to the outside. Furthermore, although the cyclic hexapeptides used as inhibitors may turn out to be transport substrates, they are apparently incapable of inducing expected conformational changes in P-gp, as noted above. It also remains to be seen whether these hexapeptides interact with human P-gp to the same extent that they interact with mouse P-gp, as there are considerable differences in substrate specificity between mouse and human multidrug transporters[11].

A major technical problem in this field has been the difficulty of crystallizing P-gps from humans and other mammals. This has been attributed to the extreme conformational flexibility of these transporters. That Aller *et al.*[2] discovered conditions for the crystallization of mouse *mdr1a* P-gp but not human P-gp, which is 87% identical in amino acid sequence, reinforces the mystery surrounding the challenge of crystallizing the mammalian ABC transporters. The authors' results could be related to the unstable nature of these proteins when expressed in *Pichia* yeast or to structural differences among the transporters. Nevertheless, this study gives hope that persistence will be rewarded and encourages the field to continue to seek crystals of other mammalian P-gps that diffract X-rays to higher resolution and that represent more physiologically relevant conformations.

1. Gottesman, M.M. & Ling, V. *FEBS Lett.* **580**, 998–1009 (2006).
2. Aller, S.G. *et al. Science* **323**, 1718–1722 (2009).
3. Chen, C.J. *et al. Cell* **47**, 381–389 (1986).
4. Gros, P., Croop, J. & Housman, D. *Cell* **47**, 371–380 (1986).
5. Ambudkar, S.V. *et al. Annu. Rev. Pharmacol. Toxicol.* **39**, 361–398 (1999).
6. Raviv, Y. *et al. J. Biol. Chem.* **265**, 3975–3980 (1990).
7. Bruggemann, E.P., Germann, U.A., Gottesman, M.M. & Pastan, I. *J. Biol. Chem.* **264**, 15483–15488 (1989).
8. Dey, S. *et al. Proc. Natl. Acad. Sci. USA* **94**, 10594–10599 (1997).
9. Loo, T.W. & Clarke, D.M. *Arch. Biochem. Biophys.* **476**, 51–64 (2008).
10. Dawson, R.J. & Locher, K.P. *Nature* **443**, 180–185 (2006).
11. Tang-Wai, D.F. *et al. Biochemistry* **34**, 32–39 (1995).

Monoclonal metastases



Whereas molecular heterogeneity of tumors has been well documented, less well understood are the molecular underpinnings of metastases. Using high-resolution chromosome mapping tools, Bova and colleagues show that distant metastases of prostate cancer are derived from a single tumor cell. The researchers studied the molecular signature using comparative genomic hybridization and significance analysis of microarray data of 85 metastases from 29 patients who had died of metastatic prostate cancer. Both analyses showed clustering within a single individual, supporting the notion of a clonal origin of the metastases. In contrast, tumors taken from the same anatomical site in different individuals or from different racial groups did not cluster. This finding was upheld at the single gene level: the androgen receptor gene was amplified variously in 15 of 17 patients but uniformly within the tumors of an individual patient. Whereas some copy number changes are found throughout a patient's samples (omniclonal) and others are not (subclonal), the individuals showed a distinct pattern of copy number changes ('personality') that is stable at multiple sites in a patient, which in one case involved 17 sites. This study illuminates the genomic profile of lethal prostate cancer; further work will be needed to show whether a similar process is seen in other cancers. (*Nat. Med.* **15**, 559–565, 2009) *LD*

ZFNs target endogenous plant genes

Progress in plant biotech has long been hampered by the lack of mutagenic and transgenic strategies to modify specific loci predictably. Zinc-finger nucleases (ZFNs)—fusions of specific DNA-recognition modules with an endonuclease domain—have been used for targeted engineering of several eukaryotic genomes and were previously shown to modify reporter genes in plants by inducing a double-stranded break at their target locus. Two groups now show that ZFNs can modify endogenous plant genes. Working with tobacco, Townsend *et al.* demonstrate relatively high-frequency disruption of acetolactate synthase isoforms, which confers resistance to imidazolinone and sulphonylurea herbicides. Shukla *et al.* introduce a gene conferring herbicide tolerance into the maize gene that encodes the enzyme catalyzing the final step in the synthesis of phytate, an antinutritional component of animal feed that also contributes to environmental pollution. The ZFN-modified corn plants faithfully transmitted these genetic changes to the next generation, and no off-target ZFN-induced changes to the genome were observed. Although both groups used herbicide-based screens, the frequencies of insertional inactivation should be adequate to use high-throughput sequencing to identify recombinants from a population of transformants. (*Nature*, published online April 29, 2009; doi:10.1038/nature07845 and 10.1038/nature07992) *PH*

Written by Kathy Aschheim, Laura DeFrancesco, Markus Elsner, Peter Hare & Craig Mak

Neutralizing antibody on tap

After 25 years of intensive HIV vaccine research, the successful induction of a broadly neutralizing antibody response remains elusive. Now, Johnson *et al.* show that a passive immunization strategy can protect against viral challenges, bypassing the need for elicitation of an adaptive immune response. As traditional passive immunization using antibody injection is impractical and prohibitively expensive for widespread use, Johnson *et al.* use an adeno-associated virus to deliver the gene encoding a known simian immunodeficiency virus (SIV)-neutralizing antibody to muscle cells of macaque monkeys. The myofiber cells in the muscle in turn synthesize antibodies and secrete them into the bloodstream. This approach maintains neutralizing antibody titers for at least 12 months. When challenged with SIV one month after immunization, none of the monkeys developed AIDS, and a majority showed no sign of infection. By contrast, four of six untreated monkeys died within 60 weeks. It is likely that this approach will find applications outside the HIV field, where more traditional vaccination strategies also fail to stimulate the production of neutralizing antibodies. (*Nat. Med.*, published online May 17, 2009; doi:10.1038/nm.1967) *ME*

Rational peptide specificity

Nature has evolved proteins that bind with precise affinity to specific interaction partners. In contrast, human-directed efforts to design peptides for research or therapeutic purposes have focused primarily on maximizing protein-binding affinity but not specificity. To address this problem, Grigoryan *et al.* devise a novel computational framework called CLASSY (cluster expansion and linear programming–based analysis of specificity and stability). This procedure optimizes a peptide sequence so as to maximize the thermodynamic energy gap between a desired interaction and off-target interactions. Simultaneously, it ensures that the peptide still binds its desired target as tightly as possible. CLASSY was used to design peptides that bind one of the 20 families of human bZIP transcription factors. Designed peptides were tested *in vitro* on a protein microarray for their affinities to a panel of bZIP fragments that represented all 20 families. In total, 8 bZIP families could be targeted specifically (target family bound more than all others), and 11 other families were bound with lower specificity. Going forward, CLASSY could be applied to design proteins such as zinc-finger, Src-homology 2 or PDZ domains or their ligands to engage in specific interactions. (*Nature* **458**, 859–864, 2009) *CM*

Pathways to germ cells

Cells that resemble germ cells have been generated *in vitro* from embryonic stem cells, but the process is poorly understood and differentiation efficiencies are low. To improve these methods, it will be important to unravel the molecular mechanisms of germ cell formation. A recent paper by Ohinata *et al.* takes a step in this direction. Using a serum-free and feeder-free system to culture early mouse embryos, the authors found that epiblast cells differentiate towards primordial germ cells (PGCs) in the presence of extraembryonic ectoderm and in the absence of visceral endoderm—but not vice versa. Bmp4, which is secreted by extraembryonic ectoderm, was sufficient to generate PGCs, whereas Cer1, produced by visceral endoderm, blocked this process. Another factor produced by extraembryonic ectoderm, Bmp8b, promoted differentiation to PGCs by suppressing the inhibitory effects of visceral endoderm. The *in vitro*–generated PGCs could become functional sperm after maturation of the cells *in vivo*, either by transplanting gonads reconstructed with several thousand PGCs or by injecting PGCs into the seminiferous tubules. (*Cell* **137**, 571–584, 2009) *KA*

nature biotechnology

Proteomic analysis of *S*-nitrosylation and denitrosylation by resin-assisted capture

Michael T Forrester[1,2], J Will Thompson[3], Matthew W Foster[4], Leonardo Nogueira[4], M Arthur Moseley[3] & Jonathan S Stamler[1,4]

We have modified the biotin switch assay for protein *S*-nitrosothiols (SNOs), using resin-assisted capture (SNO-RAC). Compared with existing methodologies, SNO-RAC requires fewer steps, detects high-mass *S*-nitrosylated proteins more efficiently, and facilitates identification and quantification of *S*-nitrosylated sites by mass spectrometry. When combined with iTRAQ labeling, SNO-RAC revealed that intracellular proteins may undergo rapid denitrosylation on a global scale. This methodology is readily adapted to analyzing diverse cysteine-based protein modifications, including *S*-acylation.

Nitric oxide exerts a ubiquitous influence on cellular signaling, in large part by means of *S*-nitrosylation/denitrosylation of protein cysteine residues[1]. It is also increasingly apparent that dysregulated *S*-nitrosylation may play a causal role in a spectrum of human diseases[2–4]. The biotin switch technique (BST)[5] is the most commonly used method to detect cellular *S*-nitrosylation and has greatly advanced the field[5–8]. It comprises three principal steps: blocking free thiols on cysteines by *S*-methylthiolation with a reactive thiosulfonate, converting SNOs to thiols with ascorbate, and labeling nascent thiols with biotin-HPDP (*N*-(6-(Biotinamido)hexyl)-3′-(2′-pyridyldithio)-propionamide). The degree of biotinylation (and hence *S*-nitrosylation) is determined by either anti-biotin immunoblotting or streptavidin pulldown followed by immunoblotting for the protein(s) of interest. The BST is, however, labor intensive and has relatively low throughput. Our simpler assay, SNO-RAC (**Fig. 1a** and **Supplementary Methods** online) uses a thiol-reactive resin instead of thiol-reactive biotin, thus combining the obligatory 'labeling' and 'pulldown' steps in the BST. As SNO-RAC results in a covalent disulfide linkage between the SNO site and resin, it is amenable to 'on-resin' trypsinization and peptide labeling, which subserve mass spectrometric methodologies.

Side-by-side comparison of the BST and SNO-RAC approaches using human embryonic kidney (HEK293) cells treated with *S*-nitrosocysteine (CysNO) and performed on individual (**Fig. 1b**) and total (**Fig. 1c**) SNO-proteins suggests superior sensitivity of SNO-RAC relative to the BST for proteins larger than ~100 kDa (**Fig. 1c**), perhaps due to the fewer precipitation steps required by SNO-RAC versus the BST. Proteins uniquely detected by SNO-RAC and identified by mass spectrometry (MS) are listed in **Supplementary Table 1** online. Improved sensitivity for high-mass proteins (and at least some lower mass species) was also evident in THP-1 monocytes (**Supplementary Fig. 1** online) and with the *S*-nitrosylated ryanodine receptor, which has a mass of 565 kDa (**Fig. 1d**). The BST methodology has been recently adapted to assay *S*-acylated proteins (ascorbate is substituted with hydroxylamine[9]), and an analogous RAC-based strategy (Acyl-RAC) showed robust results (**Fig. 1e**). Thus, overall, SNO-RAC appears to be more sensitive than the BST for high mass proteins, and at least as sensitive as the BST for proteins smaller than 100 kDa. Moreover, it shows general applicability in analyses of post-translational modifications of cysteines.

To further gauge the suitability of SNO-RAC for detecting endogenous SNO-proteins, we assayed *S*-nitrosylation of glyceraldehyde 3-phosphate dehydrogenase (GAPDH), a SNO-based cell-death effector[10], in RAW264.7 macrophages after stimulation of inducible nitric oxide synthase with cytokines (**Fig. 1f**). Assay specificity was confirmed by ultraviolet 'pre-photolysis' (**Fig. 1f**), which eliminates SNO but not other cysteine-based redox modifications[11]. SNO-RAC analysis of cytokine-stimulated macrophages suggested the presence of multiple *S*-nitrosylated proteins, including a novel SNO-protein at ~22 kDa, which was identified by mass spectrometry (MS) as peroxiredoxin-1 (Prx1) (**Fig. 1g** and **Supplementary Fig. 2** online). *S*-nitrosylation of Prx1 by inducible nitric oxide synthase in RAW264.7 and HEK293 cells was confirmed by SNO-RAC with western blotting (**Fig. 1h–i**).

To assess the ability of SNO-RAC to identify SNO sites, purified GAPDH treated with CysNO was compared with untreated protein after SNO-RAC, on-resin digestion with trypsin and analysis by matrix-assisted laser desorption ionization-(MALDI)-MS (**Supplementary Fig. 3** online). Four cysteine-containing peptides corresponding to GAPDH SNO-sites were identified. An analogous experiment was performed on CysNO-treated macrophages (**Supplementary Fig. 4** and **Supplementary Table 2** online), with confirmation obtained by western blotting (**Supplementary Fig. 5** online). We next attempted to characterize the full complement of *S*-nitrosylated proteins in CysNO-treated *Escherichia coli* by isotopically encoded SNO-RAC; after on-resin proteolysis, the samples were acetylated with either H_6- (control) or D_6- (CysNO-treated) acetic anhydride, which resulted in a mass shift of 3.03 Da per primary amine. This approach revealed 44 novel SNO sites in *E. coli* (**Supplementary Fig. 6** and **Supplementary Tables 3** and **4** online). Importantly, all novel identifications contained the D_3-acetylated mass, whereas an H_3-acetylated peptide was not found.

[1]Department of Biochemistry, [2]Medical Scientist Training Program, [3]Proteomics Core Facility and [4]Department of Medicine, Duke University Medical Center, Durham, North Carolina, USA. Correspondence should be addressed to J.S.S. (staml001@mc.duke.edu).

Received 7 April; accepted 6 May; published online 31 May 2009; doi:10.1038/nbt.1545

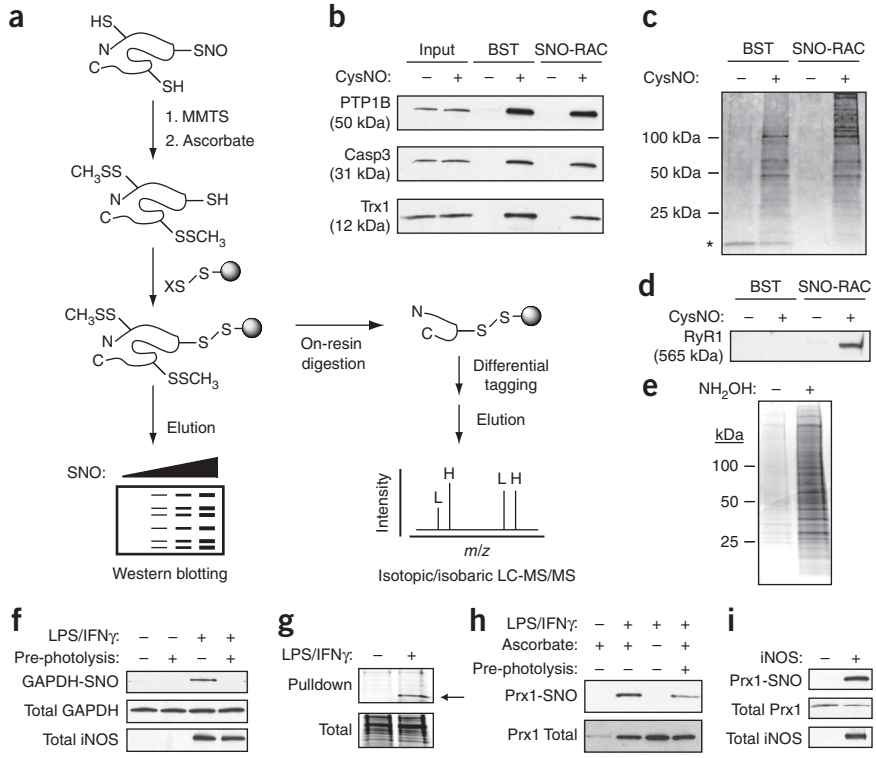


Figure 1 Analysis of protein *S*-nitrosylation by SNO-RAC. (**a**) Protein thiols are blocked with *S*-methylmethanethiosulfonate (MMTS); ascorbate converts SNOs to free thiols; nascent thiols are covalently trapped with a resin-bound 2- or 4-pyridyl disulfide (denoted X); and proteins are eluted with reductant and analyzed by SDS-PAGE. Alternatively, proteins can be trypsinized on-resin to obtain peptides containing SNO sites, which may be differentially 'tagged' with isotopic or isobaric mass labels (L, light; H, heavy), and analyzed by MS. (**b**,**c**) HEK293 cells were incubated in the presence or absence of 0.5 mM CysNO for 10 min and subjected to a side-by-side comparison of the BST and SNO-RAC approaches (1 mg protein per sample). Proteins were separated by SDS-PAGE and visualized by western blotting (**b**) or silver staining (**c**) (see **Supplementary Tables 1** and **6** online). Adventitiously eluted streptavidin is indicated by an asterisk. (**d**) Isolated rabbit sarcoplasmic reticulum (1 mg) was treated ± 0.1 mM CysNO for 10 min, subjected to either the BST or SNO-RAC; western blotting was carried out to detect the 565 kDa ryanodine receptor 1 (RyR1). (**e**) Acyl-RAC was performed on HEK293-derived membranes (1 mg protein per sample). Proteins were visualized by SDS-PAGE and Coomassie staining. (**f**) RAW264.7 murine macrophages were treated ± 0.5 μg/ml LPS and 100 U/ml IFNγ to induce inducible nitric oxide synthase, subjected to SNO-RAC (1 mg protein per sample) and analyzed as in **c** for *S*-nitrosylated GAPDH. Where indicated, extracts were photolyzed to verify assay specificity. (**g**–**i**) Macrophages were subjected to SNO-RAC (8 mg protein per sample) and eluants visualized by SDS-PAGE and Coomassie staining. The prominent band at ~22 kDa was excised and identified by MALDI-MS as peroxiredoxin-1 (indicated by arrow) (**g**). Lysates (1 mg protein per sample) from stimulated macrophages (**h**) and inducible nitric oxide synthase–transfected HEK293 cells (**i**) were subjected to SNO-RAC and analyzed for *S*-nitrosylated Prx1. For all experiments, $n \geq 3$. Full-length scans are available in **Supplementary Figure 7** online.

Enzymatic pathways for protein denitrosylation have been recently identified[12,13], but the extent to which they influence SNO-based signaling or stress is unknown. SNO-RAC was used to examine the dynamics of *S*-nitrosylation/denitrosylation in intact cells. Most SNO proteins underwent rapid denitrosylation after CysNO treatment (**Fig. 2a**). To quantify relative rates of SNO-site denitrosylation, SNO peptides derived from untreated cells or those harvested 10, 30 and 50 min after CysNO exposure, were labeled on-resin with isobaric tags for relative and absolute quantification (iTRAQ; with 114, 115, 116 and 117 amu 'reporter ions', respectively)[14], allowing for facile multiplexing with liquid chromatography–tandem MS. Kinetic analysis revealed varying rates of SNO turnover among individual sites of protein *S*-nitrosylation. A typical iTRAQ pattern is exemplified by the Cys[166]-containing ERK2 peptide,

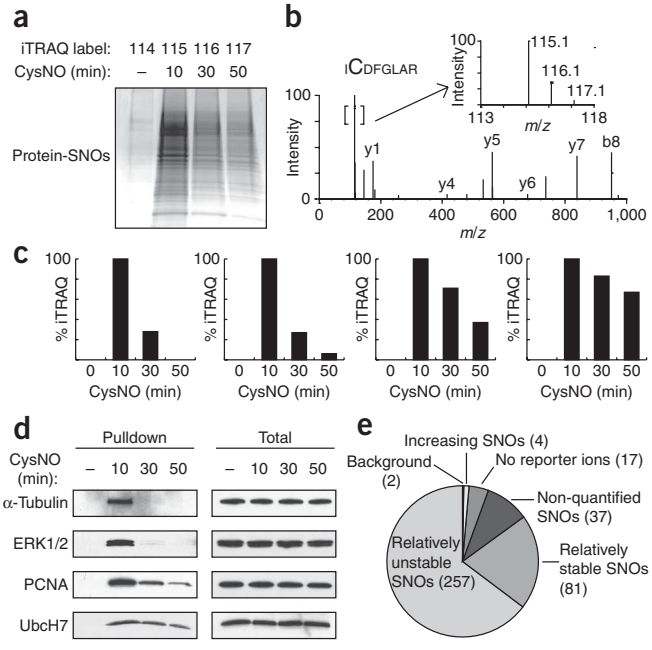


Figure 2 iTRAQ-coupled SNO-RAC demonstrates proteome-wide (global) denitrosylation. HEK293 cells were treated ± 0.5 mM CysNO for 10, 30 or 50 min and subjected to SNO-RAC. (**a**) Samples were analyzed by SDS-PAGE with Coomassie staining or (**b**–**e**) subjected to on-resin trypsinization and labeling with iTRAQ reagents, yielding peptides with reporter ions at 114, 115, 116 and 117 amu for untreated samples (0 min) and 10, 30 and 50 min post-CysNO treatment, respectively. (**b**) A representative MS/MS spectrum containing SNO-Cys[166] in Erk2. The inset shows an expanded view of the reporter ion region. (**c**) SNO sites corresponding to α-tubulin Cys[347] and Erk2 Cys[166], with a 115/116 ratio > 1.5 ('unstable'), and PCNA Cys[162] and UbcH7 Cys[86], which possess a 115/116 ratio between 0.6 and 1.5 ('stable'). (**d**) Verification of the iTRAQ data in **c** by western blotting for intact SNO proteins. (**e**) A global kinetic analysis of SNO stability via iTRAQ ratios identified 398 unique peptides, 396 of which contained a cysteine residue. The vast majority of peptides were relatively "unstable" (257 exhibited 115/116 > 1.5). Additional SNO peptides included: 81 with 115/116 between 0.6 and 1.5 (relatively 'stable'); 2 possessed barely appreciable reporter ion at 114 (115/114 ratio < 2.5) ('background'); 17 were without detectable reporter ion; 37 possessed reporter ion at 115 amu only. Four SNO sites appeared to increase over time (115/116 < 0.6). Full-length scans are available in **Supplementary Figure 7**.

which exhibited fairly rapid denitrosylation, as revealed by decreasing reporter ion intensities (**Fig. 2b**). The kinetics of *S*-nitrosylation/denitrosylation of four representative SNO sites—two relatively unstable (Cys^{347} of α-tubulin and Cys^{166} of ERK2) and two relatively stable (Cys^{162} of PCNA and Cys^{86} of UbcH7) (**Fig. 2c**)—were verified at the protein level (**Fig. 2d**). Interestingly, *S*-nitrosylated α-tubulin and ERK2 have been shown to be resistant to denitrosylation by glutathione *in vitro*[7], suggesting that multiple denitrosylases may be operative in cells[7,12,13].

Proteome-wide analysis by iTRAQ-coupled SNO-RAC confirms the global occurrence of denitrosylation (**Fig. 2e** and **Supplementary Table 5** online): nearly all SNO sites showed intense reporter ions at 115 amu, followed by progressive decay over time. Collectively, these results support the notion that denitrosylation is a major determinant of steady-state SNO levels. SNO-RAC and its modifications represent a highly versatile methodology that should facilitate the study of the diverse post-translational modifications of cysteine residues that may be designated the cysteome.

Accession numbers. PRIDE database: MS data have been deposited with accession codes 9735–9737, 9748.

Note: Supplementary information is available on the Nature Biotechnology website.

ACKNOWLEDGMENTS

We thank S. Nimkar (Applied Biosystems) for providing iTRAQ reagents, and Q. Sun, K. Ozawa, A. Hausladen and D. Hess for expert advice. This work was supported by National Institutes of Health grants U19-ES012496, P01-HL075443, HL075443 and HL059130.

AUTHOR CONTRIBUTIONS

M.T.F., M.W.F. and L.N. performed experiments and analyzed data. J.W.T. and M.A.M. acquired and analyzed mass spectrometry data. M.T.F. and J.S.S. wrote the manuscript.

COMPETING INTERESTS STATEMENT

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturebiotechnology/.

Published online at http://www.nature.com/naturebiotechnology/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Hess, D.T. *et al. Nat. Rev. Mol. Cell Biol.* **6**, 150–166 (2005).
2. Cho, D.H. *et al. Science* **324**, 102–105 (2009).
3. Lim, K.H. *et al. Nature* **452**, 646–649 (2008).
4. Foster, M.W. *et al. Trends Mol. Med.* **9**, 160–168 (2003).
5. Jaffrey, S.R. *et al. Nat. Cell Biol.* **3**, 193–197 (2001).
6. Greco, T.M. *et al. Proc. Natl. Acad. Sci. USA* **103**, 7420–7425 (2006).
7. Paige, J.S. *et al. Chem. Biol.* **15**, 1307–1316 (2008).
8. Hao, G. *et al. Proc. Natl. Acad. Sci. USA* **103**, 1012–1017 (2006).
9. Roth, A.F. *et al. Cell* **125**, 1003–1013 (2006).
10. Hara, M.R. *et al. Nat. Cell Biol.* **7**, 665–674 (2005).
11. Forrester, M.T. *et al. J. Biol. Chem.* **282**, 13977–13983 (2007).
12. Benhar, M. *et al. Science* **320**, 1050–1054 (2008).
13. Liu, L. *et al. Cell* **116**, 617–628 (2004).
14. Ross, P.L. *et al. Mol. Cell. Proteomics* **3**, 1154–1169 (2004).

nature biotechnology

Genome sequence of the recombinant protein production host *Pichia pastoris*

Kristof De Schutter[1,2,7], Yao-Cheng Lin[3,4,7], Petra Tiels[1,5,7], Annelies Van Hecke[1,5], Sascha Glinka[6], Jacqueline Weber-Lehmann[6], Pierre Rouzé[3,4], Yves Van de Peer[3,4] & Nico Callewaert[1,5]

The methylotrophic yeast *Pichia pastoris* is widely used for the production of proteins and as a model organism for studying peroxisomal biogenesis and methanol assimilation. *P. pastoris* strains capable of human-type N-glycosylation are now available, which increases the utility of this organism for biopharmaceutical production. Despite its biotechnological importance, relatively few genetic tools or engineered strains have been generated for *P. pastoris*. To facilitate progress in these areas, we present the 9.43 Mbp genomic sequence of the GS115 strain of *P. pastoris*. We also provide manually curated annotation for its 5,313 protein-coding genes.

The methylotrophic yeast *Pichia pastoris* is by far the most commonly used yeast species in the production of recombinant proteins[1] and is employed in laboratories around the world to produce proteins for basic research and medical applications. It is also an important model organism for the investigation of peroxisomal proliferation and methanol assimilation. The *P. pastoris* expression technology has been commercially available for many years. *P. pastoris* grows to high cell density, provides tightly controlled methanol-inducible transgene expression and efficiently secretes heterologous proteins in defined media. Several *P. pastoris*–produced biopharmaceuticals that are either not glycosylated (such as human serum albumin[2]) or for which glycosylation is needed only for proper folding (such as several vaccines[3]) are already on the market. An important recent breakthrough has been the development of *P. pastoris* strains with human-type N-glycosylation[4–6]. Humanized glycosylation will further increase the importance of *P. pastoris* for biopharmaceutical production; indeed, proteins produced with this system are moving into clinical development[7]. Moreover, monoclonal antibodies can be made at gram-per-liter scale in the humanized glycosylation-homogenous strains[8].

For further strain engineering, a better understanding of all aspects of the yeast's protein production machinery is needed, and a number of studies relating to *P. pastoris*'s secretory system and engineered promoters have been forthcoming[9,10]. To facilitate the investigation of *P. pastoris* and other methylotrophic yeasts, we present the 9.43 Mbp genomic sequence of the GS115 strain of *P. pastoris*.

RESULTS

Genome sequencing and assembly

Very little is known about the genomic features of *P. pastoris*. The *P. pastoris* genome has been shown to be organized in four chromosomes with a total estimated size of 9.7 Mbp by pulsed-field gel electrophoresis[11]. In addition they assigned 13 *P. pastoris* genes to the different chromosomes. The absence of a genetic map makes chromosome assembly a challenging task, which we completed according to the strategy outlined in **Figure 1a**. We made use of 454/Roche sequencing[12] (GS-FLX version) to highly oversample the genome (20× coverage) and generated 70,500 paired-end sequence tags, to enable the assembly of all but seven contigs into nine 'supercontigs' (plus the mitochondrial genome) using automated shotgun assembly and BLASTN-based contig end-joining (Online Methods and **Supplementary Fig. 1** online). Upon assigning these (super)contigs to the four chromosomes (Online Methods and **Supplementary Fig. 2** online), the order of the supercontigs was determined through PCR and Sanger sequencing of the amplification products. These finishing experiments allowed the reconstruction of the four chromosomal sequences (**Fig. 1b** and **Table 1**), with only two gaps remaining (one each on chromosomes 1 and 4). A ribosomal DNA (rDNA) repeat sequence was present in the assembly as a separate contig of 7,450 bp, with exceptionally high coverage (328.8-fold). Given that sequence coverage all over our assembly very closely approximates 20×, we interpret that there are ~16 copies of the rDNA repeat region, thus accounting for about 119 kbp in sequence. We detected these rDNA loci on all chromosomes (Online Methods, **Fig. 1b** and **Supplementary Fig. 2**). The rDNA locus contains the 18S, 5.8S and 26S rRNA coding sequences. Unlike the *Saccharomyces cerevisiae* 5S rRNA gene, which is localized to the repeated rDNA locus, the 21 copies of the *P. pastoris* 5S rRNA are spread across the entire length of all chromosomes. Based on pulsed-field gel electrophoresis (PFGE), the chromosomes of *P. pastoris* GS115 were estimated to be 2.9, 2.6, 2.3 and 1.9 Mbp[11], whereas we obtained 2.88 (2.8 + 0.08), 2.39, 2.24

[1]Unit for Molecular Glycobiology, Department for Molecular Biomedical Research, VIB, Ghent-Zwijnaarde, Belgium. [2]Department for Biomedical Molecular Biology, Ghent University, Ghent- Zwijnaarde, Belgium. [3]Department of Plant Systems Biology, VIB, Ghent-Zwijnaarde, Belgium. [4]Department of Plant Biotechnology and Genetics, Ghent University, Ghent, Belgium. [5]Unit for Molecular Glycobiology, L-ProBE, Department of Biochemistry and Microbiology, Ghent University, Ghent-Zwijnaarde, Belgium. [6]Eurofins MWG Operon, Ebersberg, Germany. [7]These authors contributed equally to this work. Correspondence should be addressed to N.C. (Nico.Callewaert@dmbr.vib-UGent.be).

Received 1 April; accepted 6 May; published online 24 May 2009; doi:10.1038/nbt.1544

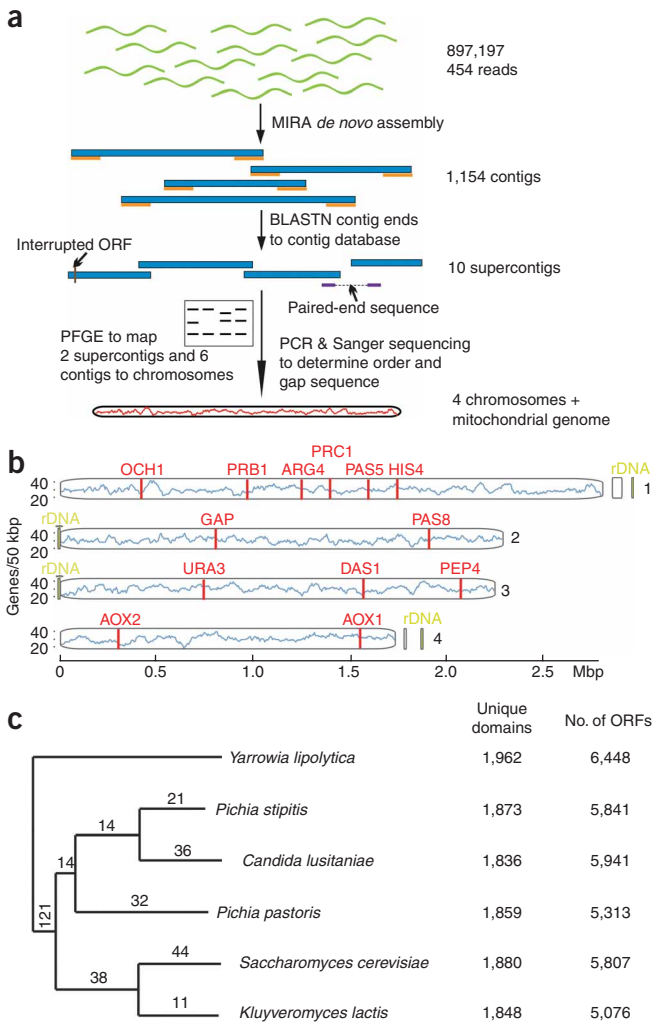


Figure 1 *Pichia pastoris* genome sequencing and overview. (**a**) Genome sequencing and assembly strategy. (**b**) *P. pastoris* gene density and known markers position. Gene density is plotted as a histogram, showing a uniform distribution of genes across each chromosome. The gene density is calculated in a window size of 50 kbp with 5 kbp sliding window. Genes that had been previously mapped to the chromosomes through PFGE are indicated in red, and rDNA repeats in green. (**c**) Phylogenetic tree. The phylogenetic tree was built on the concatenated sequence of 200 single-copy orthologous genes in all of the six species. Numbers next to each branch correspond to the number of Pfam domains uniquely present in the corresponding lineage.

and 1.8 (1.78 + 0.017) Mbp after assembly (assembled chromosome + assigned contig). Including the estimated 0.12 Mbp of rRNA repeats, we calculate a genome size of 9.43 Mbp.

Genome sequence accuracy estimation

A concern with genome sequences largely generated through 454 sequencing is the potential for 'indel errors' at homopolymeric sequences[13]. An analysis of the occurrence of such sequences in the *P. pastoris* genome is provided in **Supplementary Figure 3** online. Two approaches were followed to estimate the accuracy of our genome sequence. First, we retrieved 39 peer-reviewed Genbank coding sequences of *P. pastoris* strain GS115 (**Supplementary Table 1** online; total sequence length 70,295 bp). These sequences were compared to our genome sequence, and 84 differences were encountered. To establish which sequence was correct, we amplified these genes by PCR and Sanger-sequenced the PCR products. In all but two cases, the Sanger sequences confirmed our genome sequence, and we thus estimate the error rate to be 1 in 35,147 bp. In an alternative approach, we analyzed all open reading frames (ORFs) encoding proteins with at least one clear homolog in the databases. Where we found an interrupted ORF with clear homology to the 5′ part of the homologs, immediately followed by a coding sequence with clear homology to the 3′ part, the most logical interpretation was that there was a frameshift error mutation in our genome sequence (that is, both coding sequences are extremely likely to be linked into one open reading frame (ORF)). We found such frameshift errors in 2.7% (108) of the 3,997 genes for which such analysis could be made, totaling 6.11 Mbp of coding sequence. Conservatively estimating that we would only have detected such error if it occurred in the first two-thirds of the ORF, we then calculated a frameshift error rate in the coding sequences of 1 in 37,716 bp. Both estimates show that high-coverage 454 sequencing can indeed yield highly accurate genome sequences.

Pichia pastoris phylogenetic position

Phylogenetic analysis (**Fig. 1c**; Online Methods) shows that *P. pastoris* diverged before the formation of the CTG clade (yeasts which translate the CUG codon into serine instead of leucine[14]).

Genome sequence annotation: protein-coding genes

Protein-coding genes were automatically predicted using EuGène[15] (Online Methods and **Supplementary Fig. 4** online). The gene models were manually curated for functional annotation, accurate translational start-and-stop assignment, and intron location. This resulted in a 5,313 protein-coding gene set of which 3,997 (75.2%) have at least one homolog in the National Center for Biotechnology Information protein database (BLASTP e-value 1e-5, sequence length $\leq$20% difference and sequence similarity $\geq$50%). The protein-coding genes occupy 80% of the genome sequence. According to recently proposed measures for genome completeness, we searched the genome for highly conserved single (or low) copy gene sets: core eukaryotic genes (CEGs) with 248 genes across six model organisms[16] and FUNYBASE[17] with 246 genes with orthologs in 21 fungi. All genes from both gene sets were present in our proteome with full domain coverage.

We assigned 1,285 genes to the Kyoto Encyclopedia of Genes and Genomes (KEGG) metabolic pathways, and 4,262 of the genes were annotated with Gene Ontology (GO) terms[18]. The GO slim categories of *P. pastoris* are presented in **Supplementary Figure 5** online. A secretion signal peptide was predicted in 9% of the genes[19], and 4,274 of proteins contain InterPro domains. These include 2,320 distinct Pfam domains. In comparing the presence and absence of protein domains with five other yeast proteomes, 32 domains in 32 genes are identified as specific to *P. pastoris* (**Supplementary Table 2** online). The two fungi in the CTG clade whose genomes have been sequenced (*P. stipitis* and *C. lusitaniae*) share 71 gene families that are absent in *P. pastoris* (**Supplementary Table 2**).

Codon (pair) optimization of transgenes to the expression host organism often yields substantial improvements in recombinant protein yield[20]. *P. pastoris*'s codon usage is shown in **Figure 2a**, which will guide synthetic gene design for protein production in this organism. Overall, the codon usage is similar to the one for *S. cerevisiae*. Some synonymous codon pairs are also more or less frequently used than expected (the codon pair bias)[21]. As reported for

Table 1 Genome sequencing and assembly statistics and contents overview

a. Genome sequencing and assembly statistics

454 Sequencing						
Sequenced reads		Sequenced length (bp)		Paired-end reads		
897,197		218,602,026		11,538		
MIRA assembly						
Assembled reads	Assembled contigs	Contigs (>500 bp)	Length (bp)	N50	L50	Average coverage
885,659	1,154	230	9,658,092	40	77	20
Contig joining			**Chromosomes**			
Joined contigs	Supercontigs	Length (Mbp)				
203	10	9.3	4			

b. Genome contents overview

General information	Coding genes	RNA genes	Mitochondrial genome
Size (Mbp): 9.3 (not including rDNA loci, estimated at 0.12 Mbp)	Coding genes: 5,313	tRNA genes: 123	Size (bp): 36, 119
Genome GC content (%): 41.1	Coding (%): 79.6	5S rRNA genes: 21	Genome GC content (%): 22
Assembled chromosomes: 4	Coding GC (%): 41.6		Coding genes: 16
	Mean gene length (bp): 1,442		tRNA genes: 31
	Single exon genes: 4,680		

N50, number of contigs that collectively cover at least 50% of the assembly. L50, length of the shortest contig among those that collectively cover 50% of the assembly.

S. cerevisiae[22], under-represented and over-represented codon pair clusters were observed (**Fig. 2b**). It remains untested in *P. pastoris* whether optimizing genes to this codon pair bias results in higher protein expression levels.

Genome sequence annotation: tRNA genes

tRNA coding genes were automatically predicted and manually confirmed by BLASTN with *S. cerevisiae* homologs, which identified 123 nuclear tRNA genes (**Supplementary Table 3** online), compared

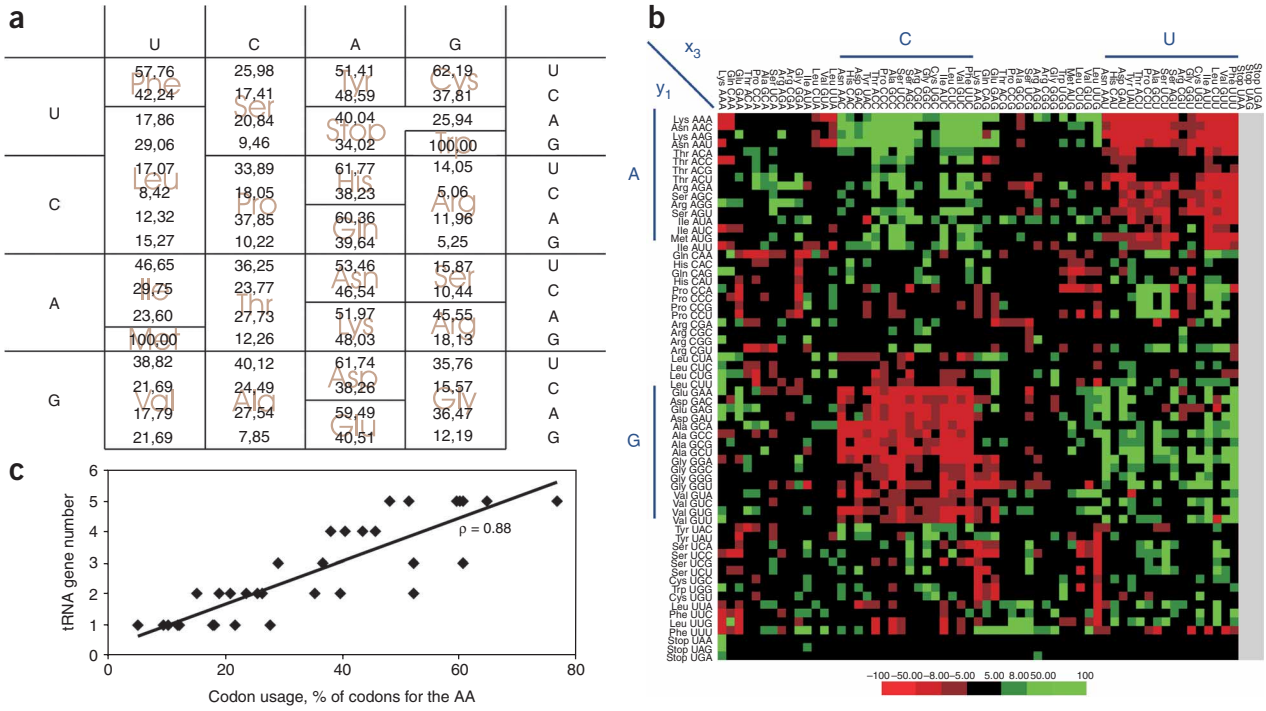


Figure 2 *Pichia pastoris* codon usage. (**a**) Codon usage. Codon usage in the *P. pastoris* ORFeome. The relative abundance of a codon is represented as a percentage of the total codon usage for the amino acid. (**b**) Codon pair usage. Codon pair residual values for *P. pastoris*. The horizontal and vertical axis show, respectively, the 5′ P-site and 3′ A-site codon. Each pixel represents a codon pair residual value. Favored codon pairs are represented in green, under-represented pairs in red. Grouping codon pairs by the x_3 and y_1 nucleotides in the $x_1x_2x_3$ and $y_1y_2y_3$ codon pair reveals over- and under-represented clusters. (**c**) Correlation of tRNA genes and codon usage. Graph shows correlation between the codon usage in relation to the number of genes coding for tRNAs recognizing this codon (Spearman $\rho = 0.88$, $P < 0.0001$).

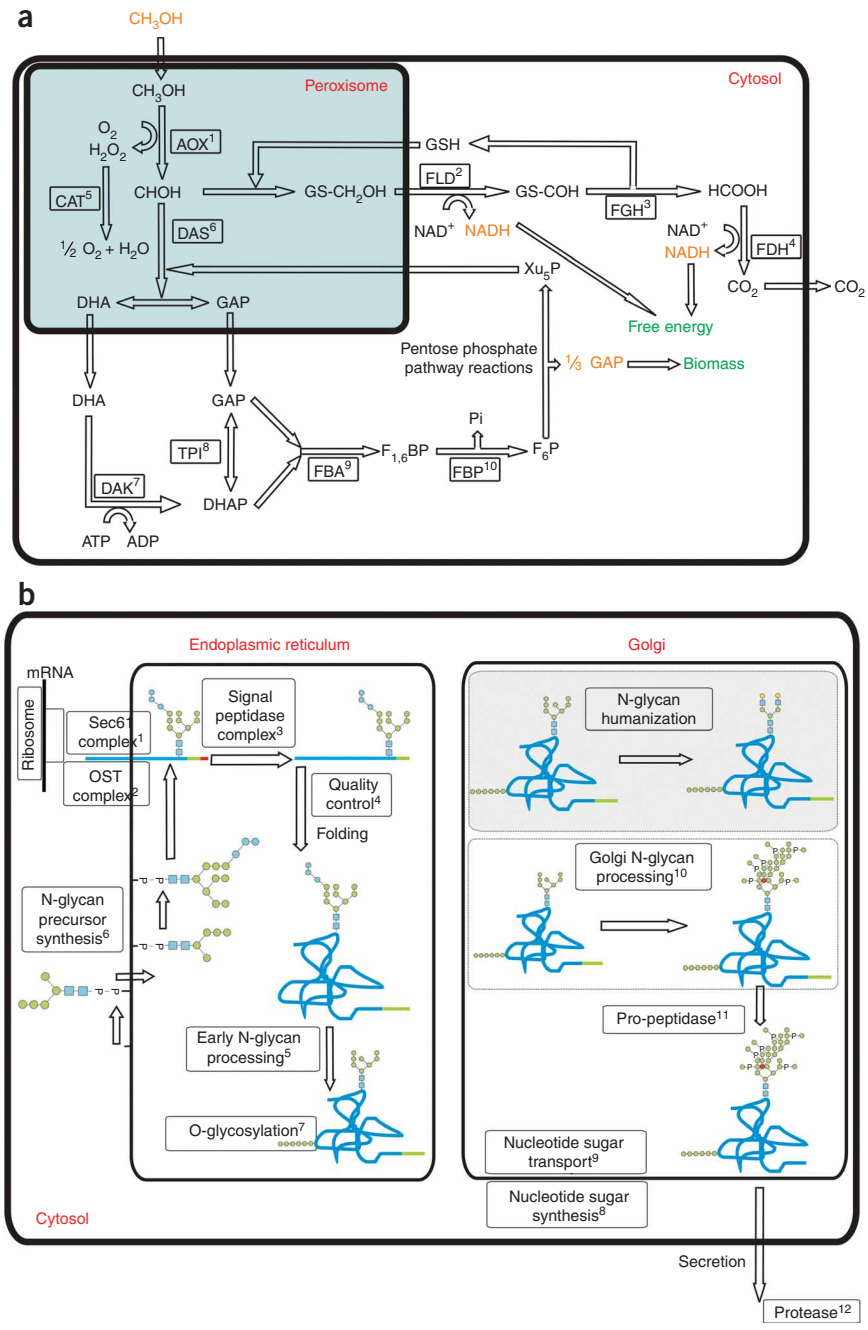


Figure 3 *Pichia pastoris* pathways. (**a**) Methanol utilization pathway in *Pichia pastoris*. A detailed table with the genes coding for the respective enzymes is shown in **Supplementary Table 4a**. [1]AOX, alcohol oxidase; [2]FLD, formaldehyde dehydrogenase; [3]FGH, S-formylglutathione hydrolase; [4]FDH, formate dehydrogenase; [5]CAT, catalase; [6]DAS, dihydroxyacetone synthase; [7]DAK, dihydroxyacetone kinase; [8]TPI, triosephosphate isomerase; [9]FBA, fructose-1,6-bisphosphate aldolase; [10]FBP, fructose-1,6-bisphosphatase; DHA, dihydroxyacetone; GAP, glyceraldehyde-3-phosphate; DHAP, dihydroxyacetone phosphate; $F_{1,6}BP$, fructose-1,6-bisphosphate; F_6P, fructose-6-phosphate; P_i, phosphate; Xu_5P, xylulose-5-phosphate; GSH, glutathione. (**b**) Protein secretion pathway. Schematic representation of the secretion pathway in *P. pastoris*. A detailed table with the genes coding for the components involved in the represented complexes or processes is shown in **Supplementary Table 4b**. The nascent protein is translocated to the ER by the Sec61 complex, and N-glycosylation sites are glycosylated with the dolichol-linked $Glc_3Man_9GlcNAc_2$ oligosaccharide precursor by the OST complex. After processing of the signal peptide, the protein is folded with the aid of chaperones. ER N-glycan processing results in $Man_8GlcNAc_2$ type glycan. O-glycosylation is also initiated in the ER by the protein-*O*-mannosyltransferases. After transport to the Golgi apparatus, the N-glycans are further processed to the yeast-typical hypermannosyl-type glycans. In strains with humanized glycosylation pathways,[4–6] the hypermannosylation is abolished and the glycans are processed to $Gal_2GlcNAc_2Man_3GlcNAc_2$. After processing of the pro-domain, the protein is secreted in the growth medium, where it may be a substrate for yeast proteases.

to 274 in the *S. cerevisiae* genome[23]. *P. pastoris* has three tRNA families not present in *S. cerevisiae* (tR(UCG), tL(CAG) and tP(CGG)), but also lacks one tRNA family (tL(GAG)).

Notably, a positive correlation was found between the number of tRNA genes for a given codon and the frequency of use of this codon (Spearman $\rho = 0.88$; $P < 0.0001$, **Fig. 2c**).

DISCUSSION

The genomic sequence of *P. pastoris* presented here will facilitate the development of improved strains with customized properties for high-yield protein production with defined post-translational modifications. Promising targets for genetic engineering include inducible promoters for transgene expression, chaperones that assist protein folding, proteins involved in the secretory pathway and enzymes catalyzing protein glycosylation, proteolytic processing and other post-translational modifications.

The commonly used methanol-inducible promoters in *P. pastoris*—the alcohol oxidase I promoter[10,24] and the formaldehyde dehydrogenase promoter[25]—drive the production of enzymes needed for methanol assimilation and therefore produce extremely high levels of these transcripts upon switching the carbon source to methanol. The genome sequence has allowed identification of all genes coding for enzymes involved in methanol assimilation and their promoters (**Fig. 3a** and **Supplementary Table 4a** online), which can now be studied for their suitability for transgene expression in *P. pastoris*. A first comparative analysis of these promoters did not reveal obvious commonalities in sequence motifs or promoter organization (data not shown).

Secretion of heterologous proteins rather than cytoplasmic accumulation is most often the preferred option in *Pichia*-based production processes. The yeast secretory system (overview in **Fig. 3b**; **Supplementary Table 4b** summarizes the genes discussed in the remainder of the text) is thus an important engineering target to obtain optimized strains that are capable of folding and processing a large flux of recombinant protein. However, many aspects of the secretory pathway are insufficiently characterized. For example, the

knowledge on the *Pichia* chaperones is incomplete, and we here provide the complete catalog of orthologs to the *S. cerevisiae* endoplasmic reticulum (ER) folding machinery, which should enable more efficacious folding-system engineering in the future[26].

The heterologous preprot signal sequence of the *S. cerevisiae* alpha-mating factor is most often used to induce Sec61p-mediated translocation of the protein into the endoplasmic reticulum of *P. pastoris* (http://faculty.kgi.edu/cregg/). This signal sequence works in most cases, although there have been almost no studies to compare it to other signal sequences. Moreover, the Kex2p/Ste13p-mediated processing of the propeptide in this *S. cerevisiae* sequence is often problematic in *Pichia*[27], resulting in nonnative amino acids at the N-terminus of the heterologous protein. The genome sequence now reveals a multitude of endogenous signal sequences (**Supplementary Fig. 6** online shows a subset of such signal sequences, derived from homologs of functionally annotated secreted *S. cerevisiae* proteins). This database of secretion signals will allow screening for the optimal signal-ORF combination, which may result in augmented protein expression levels. Multiple sequence alignment also allowed derivation of a consensus signal sequence (**Supplementary Fig. 6**), which may be suited for mediating heterologous protein secretion.

The secretory system is also the site of post-translational modification (especially glycosylation), and yeasts differ substantially from higher eukaryotes in this respect. In terms of N-glycosylation, yeasts such as *P. pastoris* modify proteins with a range of heterogenous high-mannose glycans[28], which introduce a large amount of heterogeneity in the protein (reducing downstream processing efficiency and complicating product characterization) and induce fast clearance from the bloodstream. The highly immunogenic terminal alpha-1, 3-mannosyl glycotopes that are abundantly produced by *S. cerevisiae* are not detected on *Pichia*-produced glycoproteins[29]. Indeed, we did not find an ortholog of the *S. cerevisiae* gene *MNN1* (encoding the alpha-1,3-mannosyltransferase) in the *Pichia* genome. However, *Pichia* glycoproteins can in some cases be modified with β-1,2-mannose residues[30], reminiscent of antigenic epitopes on the *Candida albicans* cell wall[31]. We find the patented *P. pastoris* AMR2 β-mannosyltransferase in the genome, and three homologs, thus providing the basis for reducing the levels of this undesired glycan modification.

To overcome the difficulties with *Pichia*'s glycosylation, strains have been developed with an entirely re-engineered glycosylation pathway to produce human IgG–type N-glycans (N-glycosylation humanization technology; **Fig. 3b**)[4–6]. The heterologous glycosyltransferases needed for this use the sugar-nucleotides UDP-GlcNAc and UDP-Gal as monosaccharide donors. Although UDP-GlcNAc is synthesized in yeasts for the synthesis of cell wall chitin (we have identified a UDP-GlcNAc transporter in the genome), no galactosylated glycoconjugates in *P. pastoris* have been described. We have shown previously that the mere overexpression of a *Pichia* Golgi-targeted version of human β-1,4-galactosyltransferase I is sufficient to achieve galactosylation of secreted glycoproteins, indicating that *Pichia* produces UDP-Gal and transports it into the Golgi apparatus[32]. Indeed, we now find an endogenous cytoplasmic UDP-Glc-4-epimerase and clear homologs of Golgi UDP-Galactose transporters in the *P. pastoris* genome (**Supplementary Table 4b**). These findings are relevant to glycan engineering in this yeast as researchers have previously overexpressed a heterologous UDP-Glc-4-epimerase in fusion to the galactosyltransferase to achieve higher levels of UDP-Gal in the yeast Golgi apparatus[6,33].

Yeasts also O-glycosylate secreted proteins with oligomannosyl-glycans that differ from the mucin-type O-glycosylation in humans[34]. No robust engineering approach has yet been developed to overcome this issue. The identification of the *Pichia* protein-O-mannosyltransferases that initiate this modification in the ER in the genome will help toward this goal.

Finally, an often-observed problem is degradation of the product by endogenous proteases. If the heterologous protein is toxic to the cell, much of this proteolytic activity can be of vacuolar origin (released in the growth medium upon cell lysis), but *Pichia* also expresses secreted proteases. It would be of great interest to have a panel of *P. pastoris* strains in which the most active proteases had been disrupted. Only few such strains are currently available because knowledge on the protease gene sequences was unavailable. We here provide a catalog of the *Pichia* vacuolar and secreted proteases (**Supplementary Table 4b**), which will speed up the development of protease-deficient strains.

The wealth of information provided by a full genome sequence will enable a more rapid development of *P. pastoris* as a protein expression host, building on its exceptional natural capacity for heterologous protein production. With a large academic and industrial user base, human-type N-glycosylation already in place, gram-per liter monoclonal antibody production recently reported[8] and the genome now publicly available, the stage is set for *Pichia pastoris* to become an even more important expression system for biopharmaceutical proteins.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturebiotechnology/.

Accession numbers. The *P. pastoris* genomic sequence has been deposited in the EMBL Nucleotide Sequence Database (Accession numbers FN392319–FN392325).

Note: Supplementary information is available on the Nature Biotechnology website.

ACKNOWLEDGMENTS

This research was supported by a Marie Curie Excellence Grant to N.C. (EU-FP6), IUAP P6/25 (BioMaGNet) and by the Fund for Scientific Research-Flanders (FWO). Y.-C.L. is supported by the EVOLTREE (EU-FP6) fellowship. Contributions towards the sequencing cost were received from VIB and from Research Corporation Technologies. We thank Lieven Sterck and Kenny Billiau for setup and maintenance of the BOGAS annotation portal and Cindy Martens for the Dollo parsimony analysis. We thank Mark Veugelers and Jo Bury for continuous support of this project.

AUTHOR CONTRIBUTIONS

K.D.S. and P.T. assembled and finished the genome sequence, manually curated the computer-generated annotation, analyzed the annotation and wrote parts of the manuscript. Y.-C.L. performed all post-shotgun assembly bio-informatics aspects of the study under guidance of Y.V.d.P. and P.R. and wrote parts of the manuscript. A.V.H. assisted in gap closure and in determining sequence accuracy. S.G. performed the 454/Roche sequencing and J.W.-L. processed the raw data and performed shotgun assembly and contig scaffolding. Both provided the corresponding methods sections of the manuscript. N.C. designed and coordinated the study, initiated the BLAST-based contig joining approach and wrote parts of the manuscript.

Published online at http://www.nature.com/naturebiotechnology/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Cregg, J.M., Cereghino, J.L., Shi, J. & Higgins, D.R. Recombinant protein expression in *Pichia pastoris*. *Mol. Biotechnol.* **16**, 23–52 (2000).
2. Watanabe, H. *et al.* In vitro and in vivo properties of recombinant human serum albumin from *Pichia pastoris* purified by a method of short processing time. *Pharm. Res.* **18**, 1775–1781 (2001).
3. Hardy, E. *et al.* Large-scale production of recombinant hepatitis B surface antigen from *Pichia pastoris*. *J. Biotechnol.* **77**, 157–167 (2000).

4. Hamilton, S.R. *et al.* Humanization of Yeast to Produce Complex Terminally Sialylated Glycoproteins. *Science* **313**, 1441–1443 (2006).
5. Hamilton, S.R. & Gerngross, T.U. Glycosylation engineering in yeast: the advent of fully humanized yeast. *Curr. Opin. Biotechnol.* **18**, 387–392 (2007).
6. Jacobs, P.P., Geysens, S., Vervecken, W., Contreras, R. & Callewaert, N. Engineering complex-type N-glycosylation in *Pichia pastoris* using GlycoSwitch technology. *Nat. Protoc.* **4**, 58–70 (2009).
7. Ratner, M. Pharma swept up in biogenerics gold rush. *Nat. Biotechnol.* **27**, 299–301 (2009).
8. Potgieter, T.I. *et al.* Production of monoclonal antibodies by glycoengineered *Pichia pastoris*. *J. Biotechnol.* **139**, 318–325 (2009).
9. Mogelsvang, S., Gomez-Ospina, N., Soderholm, J., Glick, B.S. & Stachelin, L.A. Tomographic evidence for continuous turnover of Golgi cisternae in *Pichia pastoris*. *Mol. Biol. Cell* **14**, 2277–2291 (2003).
10. Hartner, F.S. *et al.* Promoter library designed for fine-tuned gene expression in *Pichia pastoris*. *Nucleic Acids Res.* **36**, e76 (2008).
11. Ohi, H., Okazaki, N., Uno, S., Miura, M. & Hiramatsu, R. Chromosomal DNA patterns and gene stability of *Pichia pastoris*. *Yeast* **14**, 895–903 (1998).
12. Margulies, M. *et al.* Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376–380 (2005).
13. Huse, S.M., Huber, J.A., Morrison, H.G., Sogin, M.L. & Welch, D.M. Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol.* **8**, R143 (2007).
14. Fitzpatrick, D.A., Logue, M.E., Stajich, J.E. & Butler, G. A fungal phylogeny based on 42 complete genomes derived from supertree and combined gene analysis. *BMC Evol. Biol.* **6**, 99 (2006).
15. Foissac, S. *et al.* Genome annotation in plants and fungi: EuGene as a model platform. *Current Bioinformatics* **3**, 89–97 (2008).
16. Parra, G., Bradnam, K., Ning, Z., Keane, T. & Korf, I. Assessing the gene space in draft genomes. *Nucleic Acids Res.* **37**, 289–297 (2009).
17. Marthey, S. *et al.* FUNYBASE: a FUNgal phYlogenomic dataBASE. *BMC Bioinformatics* **9**, 456 (2008).
18. Schmid, R. & Blaxter, M. annot8r: GO, EC and KEGG annotation of EST datasets. *BMC Bioinformatics* **9**, 180 (2008).
19. Emanuelsson, O., Brunak, S., von Heijne, G. & Nielsen, H. Locating proteins in the cell using TargetP, SignalP and related tools. *Nat. Protoc.* **2**, 953–971 (2007).
20. Hu, S. *et al.* Codon optimization, expression, and characterization of an internalizing anti-ErbB2 single-chain antibody in *Pichia pastoris*. *Protein Expr. Purif.* **47**, 249–257 (2006).
21. Gutman, G.A. & Hatfield, G.W. Nonrandom utilization of codon pairs in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **86**, 3699–3703 (1989).
22. Friberg, M., von Rohr, P. & Gonnet, G. Limitations of codon adaptation index and other coding DNA-based features for prediction of protein expression in *Saccharomyces cerevisiae*. *Yeast* **21**, 1083–1093 (2004).
23. Hani, J. & Feldmann, H. tRNA genes and retroelements in the yeast genome. *Nucleic Acids Res.* **26**, 689–696 (1998).
24. Tschopp, J.F., Brust, P.F., Cregg, J.M., Stillman, C.A. & Gingeras, T.R. Expression of the lacZ gene from two methanol-regulated promoters in *Pichia pastoris*. *Nucleic Acids Res.* **15**, 3859–3876 (1987).
25. Shen, S., Sulter, G., Jeffries, T.W. & Cregg, J.M. A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast *Pichia pastoris*. *Gene* **216**, 93–102 (1998).
26. Gasser, B., Sauer, M., Maurer, M., Stadlmayr, G. & Mattanovich, D. Transcriptomics-based identification of novel factors enhancing heterologous protein secretion in yeasts. *Appl. Environ. Microbiol.* **73**, 6499–6507 (2007).
27. Prabha, L., Govindappa, N., Adhikary, L., Melarkode, R. & Sastry, K. Identification of the dipeptidyl aminopeptidase responsible for N-terminal clipping of recombinant Exendin-4 precursor expressed in *Pichia pastoris*. *Protein Expr. Purif.* **64**, 155–161 (2009).
28. Grinna, L.S. & Tschopp, J.F. Size distribution and general structural features of N-linked oligosaccharides from the methylotrophic yeast, *Pichia pastoris*. *Yeast* **5**, 107–115 (1989).
29. Bretthauer, R.K. & Castellino, F.J. Glycosylation of *Pichia pastoris*-derived proteins. *Biotechnol. Appl. Biochem.* **30**, 193–200 (1999).
30. Mille, C. *et al.* Identification of a new family of genes involved in beta-1,2-mannosylation of glycans in *Pichia pastoris* and *Candida albicans*. *J. Biol. Chem.* **283**, 9724–9736 (2008).
31. Dalle, F. *et al.* Beta-1,2- and alpha-1,2-linked oligomannosides mediate adherence of *Candida albicans* blastospores to human enterocytes in vitro. *Infect. Immun.* **71**, 7061–7068 (2003).
32. Vervecken, W. *et al.* *In vivo* synthesis of mammalian-like, hybrid-type N-glycans in *Pichia pastoris*. *Appl. Environ. Microbiol.* **70**, 2639–2646 (2004).
33. Bobrowicz, P. Engineering of an artificial glycosylation pathway blocked in core oligosaccharide assembly in the yeast *Pichia pastoris*: production of complex humanized glycoproteins with terminal galactose. *Glycobiology* **14**, 757–766 (2004).
34. Trimble, R.B. *et al.* Characterization of N- and O-linked glycosylation of recombinant human bile salt-stimulated lipase secreted by *Pichia pastoris*. *Glycobiology* **14**, 265–274 (2004).

ONLINE METHODS

DNA preparation. The *P. pastoris* GS115 strain (Invitrogen) is derived from the wild-type strain NRRL-Y 11430 (Northern Regional Research Laboratories). It has a mutation in the histinol dehydrogenase gene (HIS4) and was generated by nitrosoguanidine mutagenesis at Phillips Petroleum Co[35]. It is the most frequently used *Pichia* strain for heterologous protein production.

P. pastoris genomic DNA was prepared according to a published protocol[36] with minor modifications. Instead of vortexing, the samples were shaken in a Mixer Mill (Retsch) for 2 min.

Sample preparation and sequencing with Roche/454 Genome Sequencer FLX. The shotgun library of *P. pastoris* for sequencing on the Genome Sequencer FLX (GS FLX) was prepared from 5 μg of intact genomic DNA. Based on random cleavage of the genomic DNA[12] with subsequent removal of small fragments with AMPure SPRI beads (Agencourt), the resulting single-stranded (ss) DNA library showed a fragment distribution between 300 and 900 bp with a maximum of 574 bp. The optimal amount of ssDNA library input for the emulsion PCR[12] (emPCR) was determined empirically through two small-scale titrations leading to 1.5 molecules per bead used for the large-scale approach. A total of 64 individual emulsion PCRs were performed to generate 3,974,400 DNA-carrying beads for two two-region-sized 70 × 75 PicoTiterPlates (PTP) and each region was loaded with 850,000 DNA-carrying beads. Each of the two sequencing runs was performed for a total of 100 cycles of nucleotide flows[12] (flow order TACG), and the 454 Life Sciences/Roche Diagnostics software Version 1.1.03 was used to perform the image and signal processing. The information about read flowgram (trace) data, basecalls and quality scores of all high-quality shotgun library reads was stored in a Standard Flowgram Format (SFF) file which is used by the subsequent computational analysis (see below).

Within this sequencing project, a paired end library of *P. pastoris* (strain GS115) was prepared for subsequent ordering and orienting of contigs (see computational analysis below). Six micrograms of intact genomic DNA was sheared hydrodynamically (Hydroshear, Genomic Solutions) and purified with AMPure™ SPRI beads into DNA fragments ~3 kbp in length. After methylation of EcoRI restriction sites, a biotinylated hairpin adaptor was ligated to the ends of the *P. pastoris* DNA fragments, followed by EcoRI digestion with a subsequent circularization[37]. The restriction of the circularized DNA fragments with MmeI, the subsequent ligation of paired-end adaptors and the amplification of the remaining DNA fragments resulted in a double-stranded paired-end library 130 bp in length. For the following eight individual emPCRs of the paired-end library, 1.5 molecules per bead were used to generate 339,480 DNA-carrying beads of which 280,000 were loaded onto a region of a four-region sized 70 × 75 PTP. The subsequent sequencing run with the GS FLX was performed for a total of 42 cycles of nucleotide flow (see above), and the 454 Life Sciences/Roche Diagnostics software Version 1.1.03 was used to perform the image and signal processing. The information about read flowgram (trace) data, basecalls and quality scores of all high-quality shotgun library reads was also stored in an standard flowgram format file, which is used by the subsequent computational analysis.

Computational analysis of GS FLX shotgun and paired-end reads. An automatic assembly pipeline (in-house software, Eurofins MWG Operon) was used to assemble *de novo* the generated shotgun and paired-end reads.

For *de novo* assembly of the *P. pastoris* genome sequence, a total of 897,197 good quality base-called, clipped shotgun reads with an average read length of 243 bp and a total of 70,500 good quality base-called, clipped 20 bp paired-end tag reads were used.

Within this pipeline, the information about all sequences and their quality was extracted from the SFF-file into a FASTA-file and subsequently converted into CAF format, the input format of choice of the used assembler mira (version 2.9 26×3; http://www.chevreux.org/projects_mira.html) for contig creation. The provided mate and size information (that is, forward and reverse read and the 3 kbp of length) of the paired end reads was used to scaffold the resulting contigs from the *de novo* assembly[38].

Assembly (**Fig. 1a** and **Supplementary Fig. 2**). The initial assembly contained 1,154 contigs with 9.6 Mbp sequence and 20× sequencing depth. The contig N/L50 was 40/77 kbp. Assembly of the contigs was performed manually, based on homology between the contig ends. 13 contigs were assigned to chromosomes by identification of the chromosomal markers previously described[11] (Chromosome 1: HIS4, ARG4, OCH1, PAS5, PRB1, PRC1; Chromosome 2: PAS8, GAP; Chromosome 3: DAS1, URA3, PEP4; Chromosome 4: AOX1, AOX2). Starting from these contigs, contigs with homologous contig ends were identified by BLASTN search with 500–1,000 bp of the contig ends to a database with the contig sequences. Contigs sharing homology with a *P*-value < e-20 are assumed to be linked. Pools of potentially linked contigs were assembled to supercontigs by the SeqMan assembly software (DNASTAR). The resulting contig junctions were curated by removing the low-coverage ends of either joined contig. In the cases where the BLASTN P-value was >e-50, the junction was PCR-amplified and Sanger-sequenced (primer sequences: **Supplementary Table 5** online). This resulted in ten supercontigs, with 9.1 Mbp of sequence and a remaining seven unassembled contigs. The supercontig N/L 50 was 3/1.544 Mbp.

The mitochondrial genome was also assembled and had extremely high coverage (859.9-fold), indicating the presence of ~43 mitochondrial genomes per cell in *P. pastoris* when grown on glucose as a carbon source.

Gap joining and finishing. Supercontigs were linked by mapping contigs to paired-end scaffolds ($n = 1$), and automated prediction of protein-coding sequences revealed a partial ORF at the end of a supercontig, homologous to a WD40 domain protein in other yeasts (including, *Pichia guillermondii* homolog PGUG 04385). Finding the other part of this ORF on one of the unassembled contigs allowed joining of this supercontig to one of the as-yet unassembled contigs. This was confirmed by PCR and Sanger sequencing.

Seven of the nine thus-generated supercontigs could be assigned to a specific chromosome when they contained one or more of the 13 genes for which chromosomal location had been previously established[11] (**Fig. 1b** and **Supplementary Fig. 1c**). For those two supercontigs and the six unassembled contigs where this was not the case, Southern blot analysis of pulsed-field gel electrophoresis-separated *Pichia pastoris* chromosomes (see below) was used for the assignment (**Supplementary Fig. 2**). After assignment to the chromosomes, orientation of the supercontigs and contigs on the chromosomes was determined by PCR analysis with primers on the contig ends (**Supplementary Table 5**). Gaps were PCR-amplified using primers flanking these regions (**Supplementary Table 5**) and sequenced by Sanger sequencing for finishing.

We detected rDNA repeat regions by Southern blot analysis on all four PFGE-separated chromosomes (**Supplementary Fig. 2**). The Southern signal on chromosomes 1 and 4 was as strong as those on chromosomes 2 and 3 combined. Subtelomeric location of rDNA loci is frequent in yeast genomes[39]. Because of their direct repeat character, these loci resist assembly by the current methods[40]. Through PCR, we determined the location and orientation of the rDNA locus at one end of chromosomes 2 and 3 (**Fig. 1b**). Our attempts at verification of the rDNA locus position on chromosomes 1 and 4 (still containing one gap) have so far been inconclusive.

Pulsed-field gel electrophoresis. A BioRad contour-clamped homogenous electric field CHEF DRIII system was used for PFGE. Chromosomal DNA was prepared in agarose plugs with the CHEF Genomic DNA Plug kit (BioRad) following the instructions of the manufacturer. A 0.8% agarose gel in 1× modified TBE (0.1 M Tris, 0.1 M Boric Acid, 0.2 mM EDTA) was used to separate the chromosomes. The gel was electrophoresed with a 106° angle at 14 °C at 3 V/cm for 32 h, with a switch interval of 300 s, followed by 32 h with a switch interval of 600 s and 24 h with a switch interval of 900 s (ref. 11). After separation, the chromosomes were visualized with ethidium bromide, and the different contigs were mapped onto the chromosomes by Southern blot analysis. Therefore, the gel was incubated in 0.25 M HCl for 30 min, followed by capillary alkali transfer of the DNA onto a Hybond N+ membrane (Amersham). The probes were prepared by PCR on an open reading frame. For chromosome specific probes[11], a part of the coding sequence of HIS4 (chromosome 1), GAP (chromosome 2), URA3 (chromosome 3) and AOX1 (chromosome 4) was used. The probes were random labeled with α ^{32}P dCTP, using the High Prime kit (Roche).

Automatic gene structure prediction & functional annotation. Protein-coding genes were predicted by the integrative gene prediction platform

EuGene[15] (**Supplementary Fig. 4**). A specific EuGene version was trained based on 108 manually checked *P. pastoris* genes. Documented genes from *P. stipitis* and *S. cerevisiae* were used to build *P. pastoris* orthologous gene models allowing the training of *P. pastoris*-specific Interpolated Markov Models for coding sequences and introns. Splice sites were predicted by NetAspGene[41] and gene prediction from GeneMarkHMM-ES[42] trained for *P. pastoris* and AUGUSTUS[43] (*Pichia stipitis* model) were used to provide alternative gene models for EuGene prediction. The UniProt and the fungi RefSeq protein database were searched against the supercontig sequence by BLASTX to identify the coding area. We used DeCypher-TBLASTX to search the conserved sequence area between the *P. pastoris, P. stipitis* and *Candida guilliermondii* genomes.

All predicted protein-coding genes were searched against the yeast protein database, UniProt and RefSeq fungi protein database by BLASTP. Protein domains were detected by InterProScan with various databases (BlastProDom, FPrintScan, PIR, Pfam, Smart, HMMTigr, SuperFamily, Panther and Gene3D) through the European Bioinformatics Institute Web Services SOAP-based web tools. Signal peptide and transmembrane helices were predicted by SignalP and TMHMM respectively (http://www.cbs.dtu.dk/services/). GO (Gene Ontology) terms were derived from the InterProScan result and the KEGG (Kyoto Encyclopedia for Genes and Genomes) pathway and EC (Enzyme Commission) numbers were annotated by the annot8r pipeline[18].

Expert gene structure/functional annotation. The gene structure prediction and the database search results from various databases were formatted and stored in a MySQL relational database. A multiple alignment of each protein-coding gene with the top ten best hits against the UniProt, RefSeq fungi and yeast protein database was built by MUSCLE[44]. A BOGAS (Bioinformatics Online Genome Annotation System) *P. pastoris* annotation website was setup as the workspace for expert annotators. The initial aim of BOGAS is to provide a workspace for gene structure and functional annotation. The editing of gene structure or gene function assignment is directly updated to the MySQL relational database through the web interface. All of the modification from expert annotators is traceable and reversible by the database system. Once the expert annotator modifies the gene structure and changes the translated protein product, the system will automatically trigger the update function to check the protein domain and protein database. BOGAS also provides a search function where users can search for genes by sequence similarity (BLAST), gene id, gene name or InterPro domain. Each predicted *Pichia* gene's structure and the similarity search result was visually inspected through an embedded strip-down version of Artemis[45]. The splice sites of each gene were carefully checked and compared with *S. cerevisiae* and *P. stipitis* loci. A functional description of each gene was added to the gene annotation when a closely related homologous gene was available. The result of the annotation effort is available at http://bioinformatics.psb.ugent.be/webtools/bogas/.

Estimate of the gene space completeness. Parra *et al.*[16] proposed a set of core eukaryotic genes (CEGs) to estimate the completeness of genome sequencing and assembly programs. The CEGs contains 248 genes across six model organisms (*Homo sapiens, Drosophila melanogaster, Caenorhabditis elegans, Arabidopsis thaliana, S. cerevisiae and Saccharomyces pombe*) of which ~90% are single copy in *D. melanogaster, C. elegans, S. cerevisiae* and *S. pombe*. We checked our protein-coding genes with the HMM profile from the CEGs data set by the HMMER package. All of the 248 CEGs were present in our curated gene set with full HMM domain coverage. On the other hand, FUNYBASE (FUNgal phYlogenomic dataBASE)[17] provides 246 single-copy ortholog clusters in 21 sequenced fungal genomes. We extracted these single-copy protein sequences from the FUNYBASE website and built the HMM model for each cluster. The corrected *P. pastoris* protein sequences were searched with the FUNYBASE HMM database. All of the FUNYBASE models were presented in our gene catalog with complete domain coverage.

Detection of rRNA and tRNA loci. Ribosomal RNAs were detected automatically by INFERNAL 1.0 (INFERence of RNA ALignment) against the Rfam[46] database and manually confirmed by BLASTN search with *S. cerevisiae* homologs to the *P. pastoris* genome sequence. Localization of the rDNA locus was assayed by PFGE and PCR.

Transfer RNAs were automatically predicted by tRNA Scan-s.e.m. 1.21 (ref. 47) and manually confirmed by BLASTN search with the *S. cerevisiae* homologs to the *P. pastoris* genome sequence.

Codon usage. Nucleotide sequences of the predicted *P. pastoris* ORFeome were analyzed with ANACONDA 1.5 (ref. 48). In addition to calculation of the codon use, the analysis by ANACONDA generates a codon-pair context map for the ORFeome. This map shows one colored square for each codon-pair, the first codon corresponds to rows and the second corresponds to columns in the map. Favored codon pairs are shown in green, underrepresented ones are shown in red.

Phylogenetic tree reconstruction of fungal genomes. The phylogenetic tree was based on 200 single-copy genes which were present in 12 sequenced fungal genomes. A multiple sequence alignment was constructed using the MUSCLE program and gap removal by in-house script based on the BLOSUM62 scoring matrix. The maximum likelihood tree reconstruction program TREE-PUZZLE[49] (quartet puzzling, WAG model, estimated gama distribution rate with 1000 puzzling step) was used for phylogenetic tree reconstruction. The tree was well supported by 1,000 bootstraps in each node.

Comparative analysis of gene family and protein domain. The predicted proteomes used in this study were those of six hemiascomycetes (*P. pastoris, S. cerevisiae, K. lactis, P. stipitis, C. lustianiae and Y. lipolytica*)[50,51]. In order to obtain the gene families, a similarity search of all protein sequences from the six fungi (all-against-all BLASTP, e-value 1e-10) was performed. Gene families were constructed by Markov clustering[52] based on the BLASTP result. All predicted protein sequences from the six genomes were searched against the Pfam[53] database to obtain the protein domain occurrence in each species. The protein domain loss and acquisition was counted based on the Dollo parsimony principle by the DOLLOP program from the PHYLIP package[54].

Gene annotation. Available at http://bioinformatics.psb.ugent.be/webtools/bogas/.

35. Cregg, J.M., Barringer, K.J., Hessler, A.Y. & Madden, K.R. *Pichia pastoris* as a host system for transformations. *Mol. Cell. Biol.* **5**, 3376–3385 (1985).
36. Weiss, H.M., Haase, W. & Reilander, H. Expression of an integral membrane protein, the 5HT5A receptor. *Methods Mol. Biol.* **103**, 227–239 (1998).
37. Korbel, J.O. *et al.* Paired-end mapping reveals extensive structural variation in the human genome. *Science* **318**, 420–426 (2007).
38. Pop, M., Kosack, D.S. & Salzberg, S.L. Hierarchical scaffolding with Bambus. *Genome Res.* **14**, 149–159 (2004).
39. Venema, J. & Tollervey, D. Ribosome synthesis in *Saccharomyces cerevisiae*. *Annu. Rev. Genet.* **33**, 261–311 (1999).
40. James, S.A. *et al.* Repetitive sequence variation and dynamics in the ribosomal DNA array of *Saccharomyces cerevisiae* as revealed by whole-genome resequencing. *Genome Res.* **19**, 625–635 (2009).
41. Wang, K., Ussery, D.W. & Brunak, S. Analysis and prediction of gene splice sites in four *Aspergillus* genomes. *Fungal Genet. Biol.* **46** Suppl 1, S14–S18 (2009).
42. Ter-Hovhannisyan, V., Lomsadze, A., Chernoff, Y. & Borodovsky, M. Gene prediction in novel fungal genomes using an *ab initio* algorithm with unsupervised training. *Genome Res.* **18**, 1979–1990 (2008).
43. Stanke, M. *et al.* Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics* **7**, 62 (2006).
44. Edgar, R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
45. Rutherford, K. *et al.* Artemis: sequence visualization and annotation. *Bioinformatics* **16**, 944–945 (2000).
46. Griffiths-Jones, S. *et al.* Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res.* **33**, D121–D124 (2005).
47. Lowe, T.M. & Eddy, S.R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955–964 (1997).
48. Pinheiro, M. *et al.* Statistical, computational and visualization methodologies to unveil gene primary structure features. *Methods Inf. Med.* **45**, 163–168 (2006).
49. Schmidt, H.A., Strimmer, K., Vingron, M. & von Haeseler, A. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* **18**, 502–504 (2002).
50. Rossignol, T. *et al.* CandidaDB: a multi-genome database for *Candida* species and related Saccharomycotina. *Nucleic Acids Res.* **36**, D557–D561 (2007).

doi:10.1038/nbt.1544

51. Jeffries, T. *et al.* Genome sequence of the lignocellulose-bioconverting and xylose-fermenting yeast *Pichia stipitis*. *Nat. Biotechnol.* **25**, 319–326 (2007).
52. Enright, A.J., Van Dongen, S. & Ouzounis, C.A. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res.* **30**, 1575 (2002).
53. Finn, R. *et al.* The Pfam protein families database. *Nucleic Acids Res.* **36**, D281 (2008).
54. Felsenstein, J. Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods. *Methods Enzymol.* **266**, 418–427 (1996).

nature biotechnology

Efficient siRNA delivery into primary cells by a peptide transduction domain–dsRNA binding domain fusion protein

Akiko Eguchi[1–3], Bryan R Meade[1,2], Yung-Chi Chang[4], Craig T Fredrickson[2], Karl Willert[2], Nitin Puri[5] & Steven F Dowdy[1,2]

RNA interference (RNAi) induced by short interfering RNA (siRNA) allows for discovery research and large-scale screening[1–5]; however, owing to their size and anionic charge, siRNAs do not readily enter cells[4,5]. Current approaches do not deliver siRNAs into a high percentage of primary cells without cytotoxicity. Here we report an efficient siRNA delivery approach that uses a peptide transduction domain–double-stranded RNA-binding domain (PTD-DRBD) fusion protein. DRBDs bind to siRNAs with high avidity, masking the siRNA's negative charge and allowing PTD-mediated cellular uptake. PTD-DRBD–delivered siRNA induced rapid RNAi in a large percentage of various primary and transformed cells, including T cells, human umbilical vein endothelial cells and human embryonic stem cells. We observed no cytotoxicity, minimal off-target transcriptional changes and no induction of innate immune responses. Thus, PTD-DRBD–mediated siRNA delivery allows efficient gene silencing in difficult-to-transfect primary cell types.

RNAi has become an important technology for manipulating cellular phenotypes, mapping genetic pathways and discovering therapeutic targets, and has therapeutic potential[1–5]. However, owing to their large size (~14,000 Da) and high negative charge, siRNAs do not readily enter cells[4,5]. Indeed, naked siRNAs do not enter unperturbed cells even at millimolar concentrations[4]. Approaches for enhancing delivery of siRNAs to cells have included particle formation by means of cationic lipids, cholesterol, condensing polymers, antibody-protamine fusions and liposomes[1–5]. However, these approaches perform best with adherent tumor cells and do not work well with primary cells or nonadherent cell types, thereby severely limiting the cell types amenable to discovery research and large-scale screening with RNAi. Consequently, there is a need for an siRNA delivery approach that targets the entire cell population of all primary and tumorigenic cell types, is noncytotoxic and is independent of siRNA sequence.

In developing an siRNA delivery strategy, we started with cationic PTDs (also called cell-penetrating peptides), such as TAT, 8xArg and Antp. Cationic PTDs have been shown to deliver a wide variety of cargo into primary cells and into most tissues in preclinical models, and are currently being tested in clinical trials[6]. They are rapidly taken up into cells by macropinocytosis, a form of fluid-phase uptake performed by all cells[7,8]. Direct conjugation of cationic PTDs (6–8 positive charges) to anionic siRNAs (~40 negative charges) results in charge neutralization, inactivation of the PTD, aggregation/precipitation and cytotoxicity, with limited siRNA entry into the cells[9,10].

To circumvent PTD charge neutralization, we generated a TAT-PTD fusion protein with a single double-stranded (ds)RNA-binding domain (DRBD)[11–13], which is known to bind siRNA with high avidity ($K_D \sim 10^{-9}$)[11] and thereby masks its negative charge. DRBDs are small, ~65-residue domains that specifically bind ~12–16 bp of the dsRNA backbone on 90° surface quadrants of the dsRNA helix, resulting in four DRBDs encompassing a single siRNA (**Fig. 1a**) (**Supplementary Fig. 1** online).

To determine the ability of PTD-DRBD fusion proteins to deliver siRNAs, we generated a human H1299 lung adenocarcinoma reporter cell line using destabilized green fluorescent protein (dGFP) and destabilized red fluorescent protein (dDsRed) that allowed for direct determination of the magnitude of a single-cell RNAi response and hence, the percentage of cells undergoing an RNAi response. H1299 dGFP/dDsRed reporter cells were treated with PBS (mock), PTD-DRBD plus control siRNAs or PTD-DRBD plus one of two sequence-independent GFP siRNAs (that is, two siRNAs with different sequences targeting the same mRNA; GFP1, GFP2) and analyzed by flow cytometry for GFP knockdown at 24 h (**Fig. 1b**, left panel). PTD-DRBD delivery of GFP-specific siRNAs resulted in a substantial GFP knockdown with little to no alteration of the internal DsRed control. Similar RNAi responses were induced with three additional GFP siRNAs delivered by PTD-DRBDs (data not shown). None of the controls (nonspecific control siRNAs, PTD delivery peptide only) induced an RNAi response. PTD-DRBD–mediated GFP siRNA delivery also resulted in a substantially stronger RNAi response compared with lipofection-delivered siRNAs (**Fig. 1b**, right panel). Notably,

[1]Howard Hughes Medical Institute, [2]Department of Cellular & Molecular Medicine, UCSD School of Medicine, La Jolla, California, USA. [3]Japan Society for the Promotion of Science, Tokyo, Japan. [4]Department of Pediatrics, UCSD School of Medicine, La Jolla, California, USA. [5]Life Technologies, Austin, Texas, USA. Correspondence should be addressed to S.F.D. (sdowdy@ucsd.edu).

Received 16 January; accepted 24 April; published online 17 May 2009; doi:10.1038/nbt.1541

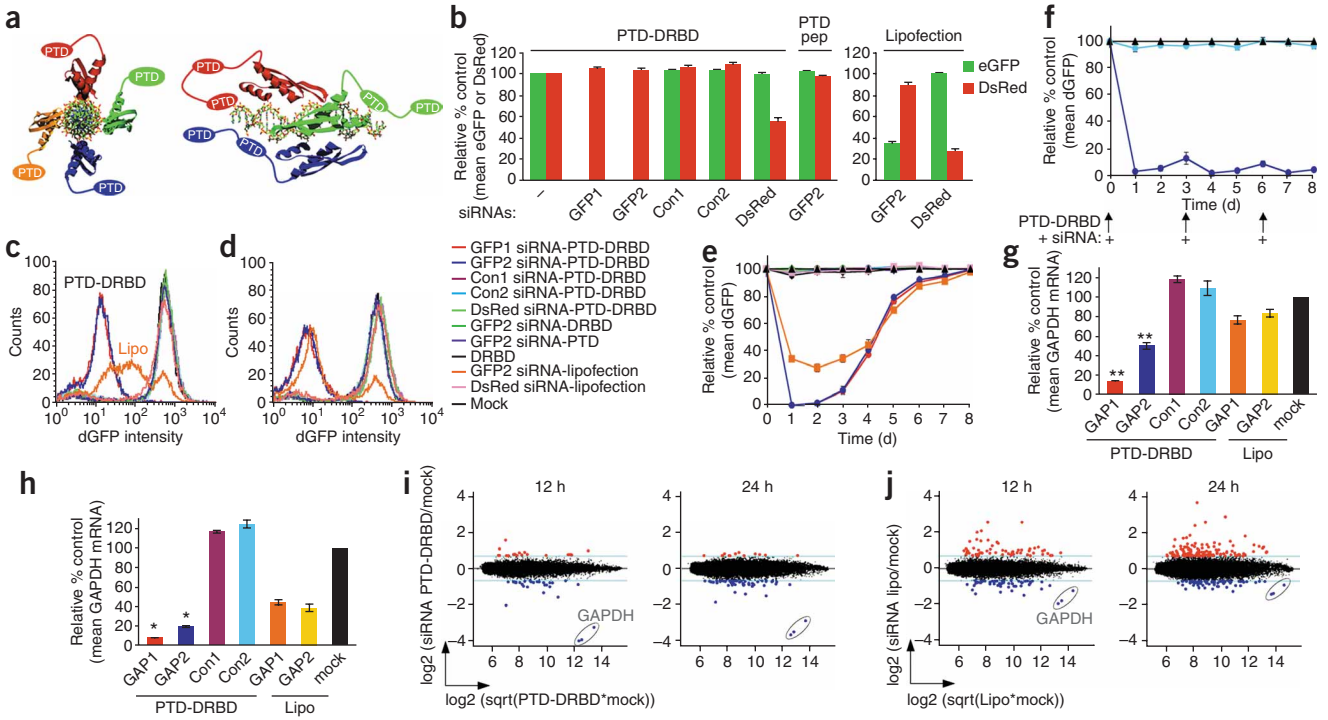


Figure 1 PTD-DRBD–mediated siRNA delivery. (**a**) Hypothetical cartoon of PTD-DRBD bound to siRNA. DRBD Ribbon structure adapted from ref. 13. (**b**) Normalized RNAi knockdown of dGFP and dDsRed by PTD-DRBD:siRNA (left panel) and lipofection (right panel), as indicated, in H1299 dGFP/dDsRed cells. Mean values were normalized to percent control. (**c**,**d**) Single-cell flow cytometry histogram analysis of dGFP RNAi response at 1 (**c**) and 2 d (**d**) after treatment of H1299 dGFP/dDsRed cells, as indicated. (**e**) Flow cytometry analysis of dGFP RNAi knockdown decay kinetics following a single siRNA treatment of dividing H1299 dGFP/dDsRed cells. Key applies to **c**,**d**,**e**. (**f**) Flow cytometry analysis of dGFP RNAi knockdown decay kinetics after multiple siRNA treatments of H1299 dGFP cells, as indicated. Mean values are normalized to percent control. (**g**,**h**) Quantitative RT-PCR analysis of endogenous GAPDH mRNA expression at 6 (**g**) and 12 h (**h**) after treatment in H1299 cells, as indicated. Mean values normalized to β2 microglobulin and reported as percent of mock-treated control. **($P < 0.001$) and *($P < 0.005$) of specific siRNA delivered by PTD-DRBD compared to lipofection. (**i**,**j**) Whole-genome microarray profile M-A plot of GAPDH siRNA delivered by PTD-DRBD (**i**) or lipofection (**j**) at 12 and 24 h after treatment in H1299 cells, as indicated. Blue line indicates 1.6-fold increase/decrease. Red and blue dots indicate genes with increased or decreased expression 1.6× background variation bar, respectively.

we detected little to no alteration of cell viability in PTD-DRBD:siRNA–treated cells, whereas lipofection resulted in varying levels of cytotoxicity (**Supplementary Fig. 2** online). Single-cell flow cytometry analysis of PTD-DRBD:GFP siRNA–treated cells showed that the entire cellular population underwent a maximal RNAi response at 24 h that was maintained at 48 h (**Fig. 1c**,**d**). In contrast, lipofection-delivered siRNAs induced a partial RNAi response, with ~20% of cells unresponsive (**Fig. 1c**,**d**). Kinetic analysis over 8 d in dividing H1299 cells showed a slow decay of the RNAi response starting 3 d after PTD-DRBD:GFP siRNA treatment that was similar to the lipofection-mediated RNAi decay kinetics (**Fig. 1e**). Similar results were obtained in primary human fibroblasts, keratinocytes, macrophages, and melanoma and glioma cells, containing integrated GFP reporter genes (**Supplementary Fig. 2**). To circumvent the RNAi decay curve, we re-treated dividing H1299 cells on days 3 and 6 with PTD-DRBD:GFP siRNAs, resulting in maintenance of the extent and magnitude of the GFP RNAi response (**Fig. 1f**).

We next targeted endogenous glyceraldehyde 3-phosphate dehydrogenase (GAPDH) mRNA by PTD-DRBD–mediated RNAi. Treatment of H1299 cells with one of two sequence-independent GAPDH siRNAs delivered by PTD-DRBD resulted in a GAPDH RNAi response that was first detected by qRT-PCR at 6 h after addition and reached a maximum by 12 h (**Fig. 1g**,**h**). In contrast, none of the PTD-DRBD negative controls induced a GAPDH RNAi response. Notably, PTD-DRBD–mediated delivery of GAPDH1 siRNA resulted in a near-maximal RNAi response by 6 h, significantly ($P < 0.001$) earlier than control lipofection delivery of the same GAPDH siRNAs (**Fig. 1g**), suggesting that PTD-DRBD–delivered siRNAs rapidly enter the cytoplasm and are loaded into the RNA-induced silencing complex. To determine whether PTD-DRBD–mediated siRNA delivery caused any cellular alterations, we examined the transcriptome of treated cells. Whole-genome microarrays were probed with total mRNA from PTD-DRBD GAPDH siRNA–treated H1299 cells at 12 h and 24 h (**Fig. 1i**). Using a 1.6-fold increase/decrease filter of cellular mRNAs, we detected a dramatic reduction in the target GAPDH mRNA along with a limited number of both up- and downregulated genes. The upregulated genes were reduced in number and the magnitude of upregulation was close to the background variations of the signal (that is, <1.6× levels) at 24 h, whereas the downregulated genes increased slightly in number and maintained a similar magnitude at 24 h (**Fig. 1i**). None of these genes are present in an innate immune response pathway or congregate into a specific genetic pathway. In contrast, lipofection-treated cells showed both a dramatic increase in both the total number of genes altered and the magnitude of the increase (**Fig. 1j**). In addition, the number of genes affected increased between 12 h and 24 h, suggesting that lipofection-based delivery substantially alters the transcriptome and may thereby confound interpretation of experimental outcomes. Lipofection induced interferon-regulated *IFIT2* and *IFIT3* genes. Moreover, lipofection-mediated GAPDH-specific knockdown was

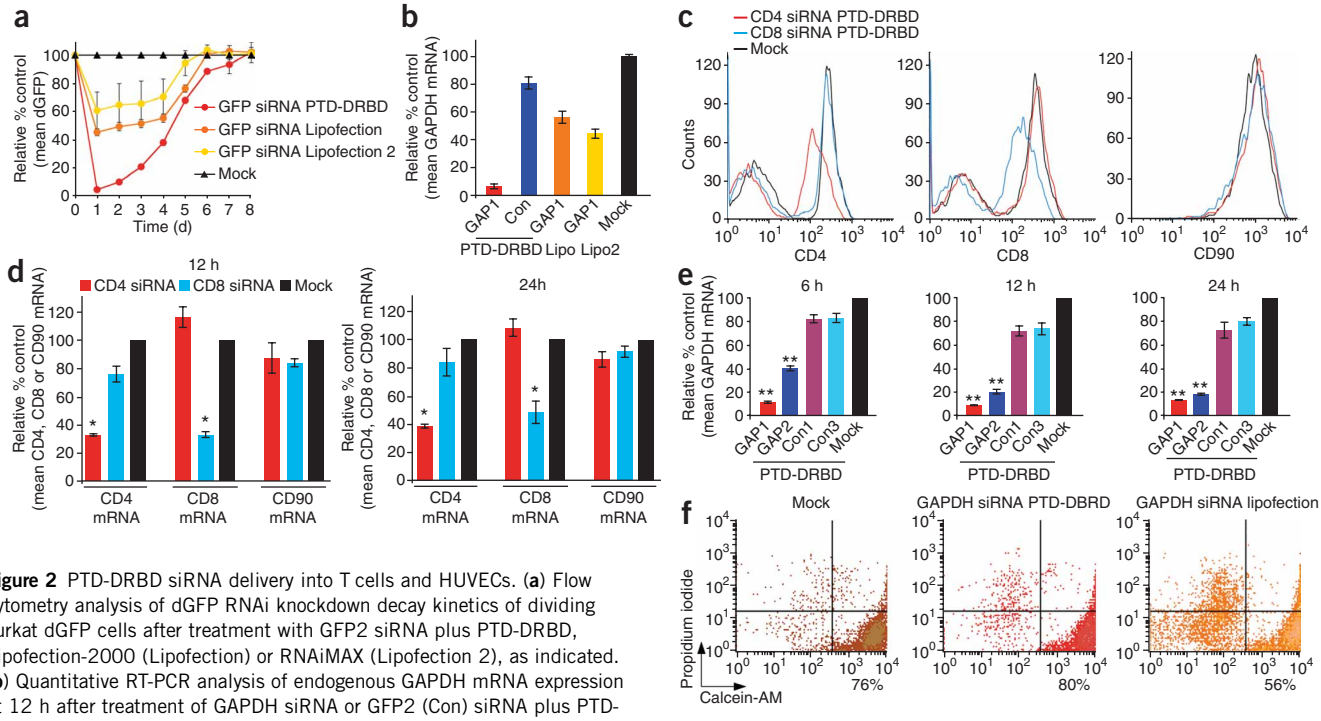


Figure 2 PTD-DRBD siRNA delivery into T cells and HUVECs. (**a**) Flow cytometry analysis of dGFP RNAi knockdown decay kinetics of dividing Jurkat dGFP cells after treatment with GFP2 siRNA plus PTD-DRBD, Lipofection-2000 (Lipofection) or RNAiMAX (Lipofection 2), as indicated. (**b**) Quantitative RT-PCR analysis of endogenous GAPDH mRNA expression at 12 h after treatment of GAPDH siRNA or GFP2 (Con) siRNA plus PTD-DRBD, GAPDH siRNA plus Lipofection-2000 (Lipo) or RNAiMAX (Lipo2) in Jurkat cells, as indicated. Con, control. Mean values normalized to β2 microglobulin and reported as percent of mock-treatment control. (**c**) Flow cytometry histogram analysis of PTD-DRBD–mediated CD4 or CD8 RNAi response at 1 d after treatment of mouse primary T cells, as indicated. (**d**) Quantitative RT-PCR analysis of endogenous CD4, CD8 or CD90 mRNA expression at 12 and 24 h after treatment of PTD-DRBD CD4 or CD8 siRNAs in primary T cells, as indicated. Mean values normalized to β2 microglobulin and reported as percent of mock-treatment control. *($P < 0.05$) of specific siRNA versus control siRNA delivered by PTD-DRBD. (**e**) Quantitative RT-PCR analysis of endogenous GAPDH mRNA expression at 6, 12 and 24 h after treatment of PTD-DRBD GAPDH or control siRNAs in primary HUVEC cells, as indicated. Mean values normalized to β2 microglobulin and reported as percent of mock-treatment control. **($P < 0.01$) of specific siRNA versus control siRNA delivered by PTD-DRBD. (**f**) PTD-DRBD cytotoxicity analysis. HUVEC cells were treated with PBS (mock treatment), GAPDH siRNA plus PTD-DRBD or lipofection and analyzed for cytotoxicity by fluorescence-activated cell sorting (FACS) after 24 h after treatment with two independent means, propidium iodide and Calcein-AM. Percent indicates viable cells present in bottom, right quadrant.

substantially smaller than PTD-DRBD–mediated knockdown. Taken together, these observations demonstrate that PTD-DRBD–mediated siRNA delivery efficiently targets the entire cellular population without cytotoxicity.

Owing to inefficient siRNA delivery and cytotoxicity, RNAi manipulation of T cells remains problematic. Therefore, we focused on a cell type that is difficult to transfect with siRNAs, namely tumorigenic Jurkat T-cells. Jurkat T-cells containing an integrated GFP reporter gene were treated with GFP siRNA plus either PTD-DRBD or one of two lipofection reagents (Lipofection-2000 and RNAiMAX) at optimal concentrations and analyzed by flow cytometry for GFP knockdown at various time points (**Fig. 2a**). PTD-DRBD delivery of GFP-specific siRNAs into Jurkat T-cells resulted in a strong GFP RNAi response in the entire population of Jurkat T-cells. In comparison, both lipofection reagents induced limited RNAi responses. Moreover, PTD-DRBD–delivered GAPDH siRNA resulted in a strong GAPDH RNAi response in Jurkat T-cells as measured by qRT-PCR, whereas the two lipofection reagents performed poorly (**Fig. 2b**).

We next treated primary murine T cells with PTD-DRBD plus CD4-specific siRNAs and assayed for cellular levels of CD4 by flow cytometry (**Fig. 2c**, left panel). The entire $CD4^+$ cellular population had undergone an RNAi response at 24 h, whereas control siRNAs did not alter CD4 levels. Similarly, PTD-DRBD–mediated delivery of CD8-specific siRNAs into primary T cells resulted in a CD8-specific RNAi response with no change in CD4 levels (**Fig. 2c**, middle panel). Consistent with these observations, we detected PTD-DRBD CD4- and CD8-specific RNAi responses by qRT-PCR at 12 and 24 h ($P < 0.05$) (**Fig. 2d**). Notably, both flow cytometry and qRT-PCR analyses of internal control CD90 receptor showed little to no alteration in T cells treated with either PTD-DRBD CD4 or CD8 siRNA (**Fig. 2c,d**). In contrast, we did not detect any RNAi responses in primary T cells using lipofection reagents (data not shown).

Primary human umbilical vein endothelial cells (HUVEC) are an important cell type for large scale RNAi screen; however, lipofection delivery of HUVECs results in both poor siRNA delivery and cytotoxicity. We targeted endogenous GAPDH mRNA by PTD-DRBD mediated RNAi. Consistent with the observations in H1299 cells above, treatment of primary HUVECs with one of the GAPDH siRNA (GAP1) delivered by PTD-DRBD resulted in a GAPDH RNAi response that reached a maximal RNAi response by 6 h ($P < 0.01$) (**Fig. 2e**), while the other (GAP2) induced significant reduction in the GAPDH mRNA after 6 h ($P < 0.01$) with a maximum after 12 h. In contrast, all PTD-DRBD negative controls failed to induce a GAPDH RNAi response. Notably, we detected little to no alteration of HUVEC cell viability in PTD-DRBD:siRNA treated cells compared to mock treated control cells (**Fig. 2f**). By contrast, siRNAs delivered into HUVECs by lipofection induced significant cytotoxicity (**Fig. 2f**).

Human embryonic stem (hES) cells have great potential to treat human disease[14]; however, manipulation of hES cells with RNAi requires protocols that target the entire cell population in a non-cytotoxic manner. We first tested the ability of PTD-DRBD to deliver siRNAs into H9 hES cells stably expressing GFP. PTD-DRBD–mediated

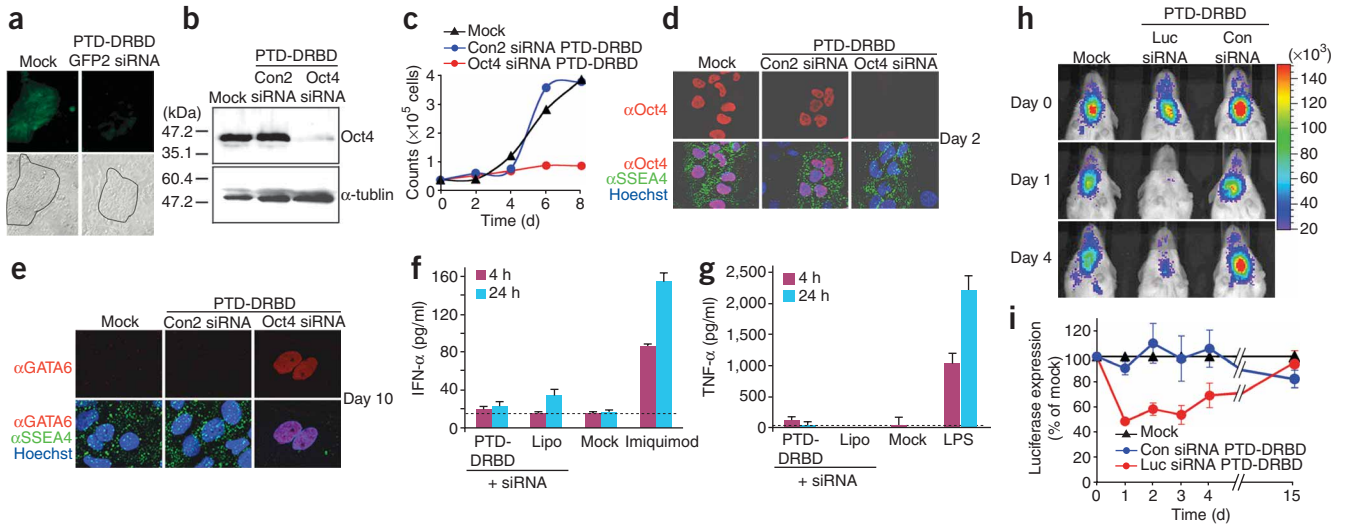


Figure 3 PTD-DRBD–mediated RNAi responses. (**a**) Fluorescent microscopy analysis of H9 hES cells constitutively expressing GFP, treated with PTD-DRBD–delivered GFP2 siRNA at 2 d after addition. Black line outlines hES cell colony on mouse feeder cell background. (**b**) OCT4 immunoblot analysis in HUES9 hES cells treated with PBS (mock treatment), PTD-DRBD–delivered OCT4 or control siRNAs at 2 d after addition. (**c**) Cell division curve of human HUES9 embryonic stem cells after mock treatment, treatment with PTD-DRBD–delivered OCT4 or control siRNAs, as indicated. (**d**) Immunohistochemistry analysis of OCT4 and SSEA4 expression in HUES9 hES cell s at 2 d after receiving mock treatment, PTD-DRB–delivered OCT4 or control siRNAs. Anti-OCT4 antibodies (red), anti-SSEA-4 antibodies (green), genomic DNA (blue). (**e**) Immunohistochemistry analysis of GATA6 and SSEA4 expression in HUES9 hES cells at 10 d after receiving mock treatment, PTD-DRB–delivered OCT4 or control siRNAs. Anti-GATA6 antibodies (red), anti-SSEA-4 antibodies (green), genomic DNA (blue). (**f**,**g**) Analysis of IFN-α and TNF-α induction in human PBMCs at 4 or 24 h after mock treatment, β-gal siRNA plus PTD-DRBD or plus lipofection, as indicated. 10 μg/ml Imiquimod or 10 μg/ml lipopolysaccharide (LPS) was used as a positive control for IFN-α or TNF-α, respectively. (**h**) Transgenic mice expressing ROSA26R-luciferase throughout the nasal and tracheal passages were imaged live on day 0. Randomized groups of luciferase-expressing mice then received mock treatment, PTD-DRBD plus Luc siRNA or control GFP (Con) siRNA and were monitored daily for luciferase expression, as indicated. Scale is in photons/s/cm^2/sr. (**i**) Graph of percent luciferase knockdown mice from **h** above. Luciferase expression was normalized to mock treatment each day. Error bar indicates s.e.m., $n = 3$ for each group with two luciferase readings performed per mouse per day.

delivery of GFP siRNAs induced a marked GFP RNAi response throughout the hES cell colony (**Fig. 3a**, circled area). We next tested the ability of PTD-DRBD–mediated siRNA delivery to affect the fate of hES cells. The OCT4 (POU5F1) transcription factor is required to maintain hES cell pluripotency, and OCT4 RNAi knockdown results in hES cell cell-cycle exit and differentiation[14]. Treatment of HUES9 hES cells with PTD-DRBD plus OCT4 siRNA resulted in both an OCT4-specific knockdown followed by a reduced growth rate and cell cycle exit indicative of pluripotency loss and initiation of differentiation (**Fig. 3b**,**c**). In contrast, neither mock treatment nor PTD-DRBD plus control siRNA altered hES cell cellular morphology, growth kinetics or OCT4 expression levels.

Pluripotent hES cells express multiple cell surface markers, including stage-specific embryonic antigen-4 (SSEA-4)[15]. During differentiation into endoderm, hES cells decrease SSEA-4 expression, stop dividing, increase in size and subsequently express the GATA6 differentiation transcription factor[16]. PTD-DRBD–delivered OCT4 siRNA resulted in loss of OCT4 expression by day 2 with continued SSEA-4 expression (**Fig. 3d**). However, by 10 d after treatment, OCT4 siRNA–treated cells had lost expression of SSEA-4 and induced expression of the endoderm-specific transcription factor GATA6 (**Fig. 3e**). In contrast, hES cells mock treated or treated with PTD-DRBD plus control siRNAs did not differentiate or alter hES cell marker expression. Similar results were obtained by simultaneous PTD-DRBD–mediated knockdown of OCT4 and Nanog, another pluripotency-associated gene (data not shown). Taken together, these observations demonstrate the ability of PTD-DRBD to deliver siRNA and rapidly induce RNAi responses in three important and difficult-to-transfect cell types: T cells, HUVECs and hES cells.

RNAi has great potential to treat human disease, including nasal delivery to treat virus infection[17]; however, *in vivo* siRNA delivery remains problematic[1–5]. To evaluate the potential of the PTD-DRBD fusion protein for *in vivo* applications, we assessed its immunogenicity in primary human peripheral blood mononuclear cells (PBMCs) and quantified the ability of PTD-DRBD to deliver siRNAs and induce an RNAi response in a reporter mouse model. siRNAs have been shown to stimulate activation of toll-like receptors 3, 7, 8 and induce innate immune responses[18]. However, PTD-DRBD–mediated delivery of immunostimulatory siRNAs did not activate interferon (IFN)-α or tumor necrosis factor (TNF)-α responses (in PBMCs) above background levels (**Fig. 3f**,**g**).

We used transgenic ROSA26 mice expressing luciferase in the nasal and tracheal passages[19] to measure the efficiency of siRNA delivery by PTD-DRBD *in vivo*. After confirming luciferase expression by live-animal imaging, the mice were randomly divided into groups (**Fig. 3h**,**i**), treated intranasally with PBS, PTD-DRBD plus Luc siRNA or control siRNA and monitored daily for 15 d for luciferase expression. PBS and PTD-DRBD control siRNA–treated mice showed no change in luciferase expression over the course of the experiment. In contrast, PTD-DRBD–delivered Luc siRNA led to an extensive reduction of luciferase expression throughout the nasal and tracheal passages at day 1, which gradually recovered by day 15 (**Fig. 3h**,**i**). These observations demonstrate the ability of PTD-DRBD–mediated siRNA delivery to induce a specific RNAi response to a quantifiable target protein in a reporter mouse model *in vivo*.

siRNA delivery has become the rate-limiting barrier to efficient cell culture and preclinical and clinical usage of siRNA therapeutics[1–5]. Although current siRNA delivery approaches have merit, they generally do not target the entire population or even a high percentage of cells, especially primary cells, and often result in some degree of cytotoxicity and alterations in cell biology. In contrast, the PTD-DRBD siRNA delivery approach described here fulfills many of the criteria for an efficient siRNA delivery system for primary cells. PTD-DRBD delivered siRNAs and induced RNAi responses in the entire population of three difficult-to-transfect primary cell types (T cells, HUVECs and hES cells) rapidly and without cytotoxicity. Because DRBDs bind to dsRNAs (siRNAs) independent of sequence composition, PTD-DRBD could in theory deliver any siRNA into cells. Lastly, the intranasal knockdown of luciferase demonstrates the *in vivo* potential of PTD-DRBD–mediated siRNA delivery; however, more *in vivo* studies are needed to ascertain the full extent of *in vivo* utility. Thus, PTD-DRBD should prove useful for basic research, target screening and potential therapeutic applications.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturebiotechnology/.

Note: Supplementary information is available on the Nature Biotechnology website.

ACKNOWLEDGMENTS

We thank V. Nizet (UCSD) for PBMCs. The hES cell line HUES9 was kindly provided by D. Melton (HHMI, Harvard University). A.E. was funded by a Japan Society for the Promotion of Science Research Fellowships for Young Scientists. This work was supported by a Specialized Center of Research grant from the Leukemia & Lymphoma Society (S.F.D.), the Elsa U. Pardee Foundation (S.F.D.), the Howard Hughes Medical Institute (S.F.D.) and the California Institute for Regenerative Medicine (S.F.D.).

AUTHOR CONTRIBUTIONS

A.E. and B.R.M. designed, purified PTD-DRBD and performed RNAi experiments. Y.-C.C performed PBMC experiments. C.T.F. performed hES cell culture. K.W. supervised hES cell culture. N.P. provided siRNAs reagents. S.F.D. supervised and analyzed data. A.E. and S.F.D. contributed to writing the manuscript, and all authors discussed and refined the manuscript.

COMPETING INTERESTS STATEMENT

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturebiotechnology/.

Published online at http://www.nature.com/naturebiotechnology/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. de Fougerolles, A., Vornlocher, H.P., Maraganore, J. & Lieberman, J. Interfering with disease: a progress report on siRNA-based therapeutics. *Nat. Rev. Drug Discov.* **6**, 443–453 (2007).
2. Kim, D.H. & Rossi, J.J. Strategies for silencing human disease using RNA interference. *Nat. Rev. Genet.* **8**, 173–184 (2007).
3. Bumcrot, D., Manoharan, M., Koteliansky, V. & Sah, D.W. RNAi therapeutics: a potential new class of pharmaceutical drugs. *Nat. Chem. Biol.* **2**, 711–719 (2006).
4. Whitehead, K.A., Langer, R. & Anderson, D.G. Knocking down barriers: advances in siRNA delivery. *Nat. Rev. Drug Discov.* **8**, 129–138 (2009).
5. Behlke, M.A. Chemical modification of siRNAs for in vivo use. *Oligonucleotides* **18**, 305–320 (2008).
6. Gump, J.M. & Dowdy, S.F. TAT transduction: the molecular mechanism and therapeutic prospects. *Trends Mol. Med.* **13**, 443–448 (2007).
7. Nakase, I. *et al.* Cellular uptake of arginine-rich peptides: roles for macropinocytosis and actin rearrangement. *Mol. Ther.* **10**, 1011–1022 (2004).
8. Wadia, J.S., Stan, R.V. & Dowdy, S.F. Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis. *Nat. Med.* **10**, 310–315 (2004).
9. Turner, J.J. *et al.* RNA targeting with peptide conjugates of oligonucleotides, siRNA and PNA. *Blood Cells Mol. Dis.* **38**, 1–7 (2007).
10. Meade, B.R. & Dowdy, S.F. Enhancing the cellular uptake of siRNA duplexes following noncovalent packaging with protein transduction domain peptides. *Adv. Drug Deliv. Rev.* **60**, 530–536 (2008).
11. Bevilacqua, P.C. & Cech, T.R. Minor-groove recognition of double-stranded RNA by the double-stranded RNA-binding domain from the RNA-activated protein kinase PKR. *Biochemistry* **35**, 9983–9994 (1996).
12. Tian, B., Bevilacqua, P.C., Diegelman-Parente, A. & Mathews, M.B. The double-stranded-RNA-binding motif: interference and much more. *Nat. Rev. Mol. Cell Biol.* **5**, 1013–1023 (2004).
13. Ryter, J.M. & Schultz, S.C. Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. *EMBO J.* **17**, 7505–7513 (1998).
14. Boyer, L.A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947–956 (2005).
15. Henderson, J.K. *et al.* Preimplantation human embryos and embryonic stem cells show comparable expression of stage-specific embryonic antigens. *Stem Cells* **20**, 329–337 (2002).
16. Hay, D.C., Sutherland, L., Clark, J. & Burdon, T. Oct-4 knockdown induces similar patterns of endoderm and trophoblast differentiation markers in human and mouse embryonic stem cells. *Stem Cells* **22**, 225–235 (2004).
17. Zhang, W. *et al.* Inhibition of respiratory syncytial virus infection with intranasal siRNA nanoparticles targeting the viral NS1 gene. *Nat. Med.* **11**, 56–62 (2005).
18. Judge, A.D. *et al.* Design of noninflammatory synthetic siRNA mediating potent gene silencing in vivo. *Mol. Ther.* **13**, 494–505 (2006).
19. Safran, M. *et al.* Mouse reporter strain for noninvasive bioluminescent imaging of cells that have undergone Cre-mediated recombination. *Mol. Imaging* **2**, 297–302 (2003).

ONLINE METHODS

PTD-DRBD fusion protein construction, design and purification. pPTD-DRBD was constructed by PCR cloning of PKR DRBD-1 into a modified pTAT vector[8] resulting in TAT-TAT-HA-TAT-DRBD-6xHis (**Supplementary Fig. 1**). The hemagglutinin (HA) epitope tag was used to follow the protein by immunoblot analysis and the 6xHis tag was used for purification over the first column, Ni-NTA. PTD-DRBD expression used BL21 codon plus (DH3) *Escherichia coli* (Stratagene); cells were transformed with pPTD-DRBD, cultured at 37 °C in Luria Bertani medium, then at 25 °C for 12 h after induction with 400 μM isopropyl-β-D-thiogalactoside. Cells were recovered by centrifugation for 5 min at 4,500*g*, sonicated in buffer A (20 mM HEPES (pH 7.5), 500 mM NaCl, 5 μg/ml aprotinin, 1 μg/ml leupeptin, 0.8 mM phenylmethylsulfonyl fluoride (PMSF)) plus 20 mM imidazole and soluble protein isolated by centrifugation for 15 min at 50,000*g*. PTD-DRBD was purified by passage over a Ni-NTA column (Qiagen), followed by loading onto a Mono-S AKAT FPLC in buffer B (50 mM HEPES (pH 7.5), 20 mM NaCl, 5% glycerol) and eluted in buffer C (buffer B plus 1.5 M NaCl). Purified PTD-DRBD was desalted (PD-10) into PBS-10% glycerol, and stored at –80 °C. EGFP-PEST (dGFP) or DsRed-PEST (dDsRed) lentiviruses were constructed using pCSC-SP-CW-EGFP-PEST or pCSC-SP-CW-DSRED[20] and pd2EGFP-N1- (destabilized GFP; BD Clontech) or pDsRed-Express-DR (destabilized DsRed; BD Clontech).

Cell culture conditions. H1299, HaCaT keratinocytes, HFF primary human fibroblasts and B16F0 melanoma cells were cultured in 10% FBS-DMEM, penicillin and streptomycin. T98G glioblastoma cells were cultured in 5% FBS-MEM and antibiotics. HUVEC cells were cultured in EGM-2 MV BulletKit (Lonza). Jurkat T-cells were cultured in 10% FBS-RPMI and antibiotics. THP-1 macrophage were grown in 10% FBS-RPMI plus 1 mM sodium pyruvate, 4.5 g/l glucose, 50 μM β-mercaptoethanol and antibiotics. Primary murine T cells were recovered from mouse spleens by MACS (Miltenyi Biotec), activated with anti-CD3ε antibody for 1 d and cultured in 10% FBS-RPMI plus 2 mM L-glutamine, 55 μM β-mercaptoethanol, 20 ng/ml IL2. The hES cell line HUES9 was kindly provided by D. Melton and H9 hES cells were obtained from WiCell. H9 hES cells were grown in 20% knockout serum-DMEM-F12 plus 55 μM β-mercaptoethanol, nonessential amino acids (NEAA), Gluta-Max, 4 ng/ml bFGF and antibiotics on murine fibroblast feeder layer. HUES9 hES cells were grown in HUES medium (10% knockout serum-DMEM plus 10% plasmonate, 55 μM β-mercaptoethanol, NEAA, Gluta-Max, 4 ng/ml bFGF, antibiotics) without murine fibroblast feeder layer in media preconditioned for 24 h on murine fibroblasts. Destabilized GFP (dGFP) and DsRed (dDsRed) proteins have ~2 h and ~12 h half-lives, respectively, substantially shorter than their wild-type parental proteins (>24 h) and therefore were used as RNAi reporter targets. dGFP and dDsRed expressing cells were generated by infection with a lentivirus expressing Vesicular Stomatitis Virus Glycoprotein (VSVG) fused to dGFP or dsRED. VSVG-dGFP and/or VSVG-dDsRed infected cells were isolated by FACS.

PTD-DRBD siRNA delivery into cells. A typical PTD-DRBD siRNA delivery reaction mixed 10 μl of 1–5 μM siRNA in water with 10 μl of 10–50 μM PTD-DRBD in PBS plus 10% glycerol plus 4 μl PBS plus 10% glycerol on ice for 30 min, diluted 1:5 in medium and added to 6×10^4 cells/well in 48-well plates for 1–6 h with final siRNA concentrations between 100–400 nM. Cells were then washed with trypsin or washed in 58 μg/ml heparin sulfate plus media for 10 min to remove extracellular PTD-DRBD:siRNA, followed by addition of fresh medium plus FBS. For primary T cells, Jurkat, Namalwa and THP-1 suspension cells, 2×10^5 cells were treated with 100–400 nM siRNA:PTD-DRBD for 1 h in medium plus 10–20% Q-serum (5 ml FBS + 1 ml Source 30Q resin (Amersham Bioscience), 30 min at RT on mixing platform, followed by 0.22 μm filtration), washed 2× with media, followed by addition of fresh complete media. For H9 and HUES9 hES cells, 6.6×10^5 cells were treated with 200–400 nM siRNA-PTD-DRBD for 1 h in serum-free medium with no feeder layer, followed by 5 h in serum-free medium on fibroblast feeder layer, then 24 h with full HUES medium plus serum. For control siRNA lipofections, cells were treated with a dose curve that yielded the highest RNAi response with 100 nM siRNA in Lipofectamine-2000 (Invitrogen) or 10–50 nM siRNA in Lipofectamine-RNAiMAX (Invitrogen) per the manufacturer's instructions. siRNAs sequences used in this study: EGFP1 (Ambion predesigned siRNA), EGFP2 (Ambion Silencer GFP), GAPDH1 (Ambion), GAPDH2 (Ambion), OCT4 (Ambion predesigned), Nanog (Ambion predesigned), Sox2 (Ambion predesigned) and Silencer Negative (control 1; Ambion); luciferase (control 2; Dharmacon), DsRed (Ambion predesigned), β-gal[17] (Dharmacon).

Immunoblotting, RT-PCR and microarrays. 6×10^4 cells/well in 48-well plates were recovered with trypsin/EDTA, whole-cell lysates were prepared in RIPA buffer (1% TritonX-100, 1% sodium deoxycholate, 40 mM Tris-HCl, 150 mM NaCl, 0.2% SDS, 5 μg/ml aprotinin, 1 μg/ml leupeptin, 0.8 mM phenylmethylsulfonyl fluoride) for 30 min on ice, clarified by centrifugation and proteins resolved by 10% SDS-PAGE. Immunoblot analyses were performed on polyvinyl difluoride membranes blocked in 4% skim milk, PBS-T (0.05% PBS, Tween20) for 1 h at 21 °C, reacted with anti-OCT4 (Santa Cruz), anti-GAPDH (Santa Cruz) and anti-α-tublin (Sigma) antibodies overnight at 4 °C, then washed and exposed to HRP conjugated anti-IgG (Santa Cruz) antibodies and detected by electrochemical luminescence (Pierce). For GAPDH mRNA TaqMan RT-PCR (Applied Biosystems), 6×10^4 dGFP-H1299 cells/well in 48-well plate were treated as described above with 400 nM GAPDH, control Silencer Negative or control luciferase siRNA and total RNA isolated at 6, 12, 24, 36, 72 and 96 h after addition. 5×10^4 HUVEC cells/well in 48-well plate were treated as described above with 400 nM GAPDH, control Silencer Negative or control GFP siRNA and total RNA isolated at 6, 12 and 24 h after addition. cDNA was synthesized using Oligo-dT and GAPDH mRNA expression was detected using TaqMan probe (Ambion) on 7300 Real time PCR system (Applied Biosystems). Mean values normalized to β2 microglobulin and reported as percent of mock GAPDH control, error bar indicates s.d., all experiments performed in triplicate. For whole-genome microarray analysis, 6×10^5 H1299 cells/well in 6-well plate were treated as described above with 400 nM GAPDH or PBS. Total RNA was isolated at 12 and 24 h after addition, and used to probe whole-genome microarrays (Illumina) at Biogem core (UCSD).

Immunohistochemistry and flow cytometry analysis. Cells were fixed with 4% paraformaldehyde for 30 min at 21 °C, permeabilized in 0.1% TritonX100-PBS for 15 min at 21 °C, blocked in 3% skim milk-PBS for 30 min at 21 °C, then reacted with anti-OCT4 (Santa Cruz), anti-SSEA4 (Santa Cruz) and anti-GATA6 (Santa Cruz) antibodies in 0.1% BSA-PBS overnight at 4 °C. Cells were washed and reacted with either Alexa488 or Alexa594 conjugated anti-IgG (Molecular Probes) for 30 min at 21 °C. DNA was counterstained with Hoechst 33342 (Molecular Probes). Cells were analyzed by confocal microscopy (Olympus Flouview). For flow cytometry, 1×10^4 dGFP- and/or dDsRed-positive cells were analyzed on a FACScan (BD Biosciences).

IFN-α and TNF-α analyses. Human peripheral blood mononuclear cells (PBMCs) were isolated from healthy donors by standard density gradient centrifugation with Ficoll-Paque PLUSTM (Amersham Biosciences) at 2,000 r.p.m. for 20 min at 20 °C. To remove platelets, PBMCs were washed 4× in 50 ml PBS, centrifuged at 1,500 r.p.m. for 8 min at 4 °C. 8×10^5 freshly isolated PBMCs were treated as described above with 100 nM β-gal siRNA[18] plus either PTD-DRBD or lipofection and seeded into 96 well-plate (4×10^5 cells/well). As a positive control, PBMCs were treated with 10 μg/ml Imiquimod for IFN-α induction and 10 μg/ml LPS for TNF-α induction. Culture supernatants were collected at 4 h and 24 h after addition, and assayed for IFN-α and TNF-α by ELISA (R&D systems).

Intranasal PTD-DRBD siRNA *in vivo* delivery. Transgenic ROSA26 loxP-Stop-loxP luciferase mice[19] (Jackson Labs) were inoculated intratracheally with 30 μl of 3 mg/ml TAT-Cre[8] to turn off luciferase gene by removal of a loxP-STOP-loxP DNA transcriptional terminator genetic element. After 3 months, D-Luciferin (150 mg/kg) was administrated intraperitoneally and luciferase expression monitored by live-animal imaging (IVIS-100, Xenogen) for 5–15 min after luciferin injection, twice daily per mouse (Day 0). After this baseline measurement, mice were randomized into groups ($n = 3$) and inoculated intranasally with 60 μl (30 μl/nostril) of PTD-DRBD plus 750 pmol luciferase siRNA or control GFP siRNA or PBS:MEM (1:1) (mock) control. Luciferase expression was monitored by IVIS imaging, twice daily per mouse each day for 15 d.

Statistical analysis. Data are expressed as mean ± s.e.m. as indicated and compared by two-tailed *t*-tests. We assign statistical significance at $P < 0.05$.

20. Miyoshi, H. *et al.* Development of a self-inactivating lentivirus vector. *J. Virol.* **72**, 8150–8157 (1998).

doi:10.1038/nbt.1541

nature biotechnology

MicroRNA-mediated species-specific attenuation of influenza A virus

Jasmine T Perez[1], Alissa M Pham[1], Maria H Lorini[2], Mark A Chua[3], John Steel[2] & Benjamin R tenOever[1–3]

Influenza A virus leads to yearly epidemics and sporadic pandemics. Present prophylactic strategies focus on egg-grown, live, attenuated influenza vaccines (LAIVs), in which attenuation is generated by conferring temperature sensitivity onto the virus. Here we describe an alternative approach to attenuating influenza A virus based on microRNA-mediated gene silencing. By incorporating nonavian microRNA response elements (MREs) into the open-reading frame of the viral nucleoprotein, we generate reassortant LAIVs for H1N1 and H5N1 that are attenuated in mice but not in eggs. MRE-based LAIVs show a greater than two-log reduction in mortality compared with control viruses lacking MREs and elicit a diverse antibody response. This approach might be combined with existing LAIVs to increase attenuation and improve vaccine safety.

Influenza A virus has the propensity to mutate and exchange segments, creating the need for annual vaccines that must be constantly updated with circulating strains identified by global monitoring[1]. Current LAIVs, such as FluMist, are temperature-sensitive reassortant viruses that contain segments derived from more than one strain[2–3]. Here we sought to develop a complementary attenuation strategy for influenza virus using microRNA (miRNA)-mediated gene silencing. miRNA-mediated viral attenuation, which relies on incorporation of miRNA target sequences into viral RNA, has recently been described for lentiviruses, picornaviruses and rhabdoviruses[4–7]. Transcript regulation by miRNAs occurs through direct binding, resulting in translational repression or cleavage[8]. miRNAs exhibit broad expression patterns, ranging from ubiquitous to tissue or lineage specific, and moderately affect global protein levels[9–11]. Although most miRNAs are evolutionarily conserved, a small percentage are species specific[12].

To determine the feasibility of miRNA-mediated attenuation of influenza A virus, we investigated whether infection affects pre-miRNA formation, maturation or post-transcriptional gene silencing. Tissue-specific and ubiquitous miRNAs, miR-124 and miR-93, respectively, were expressed exogenously in HEK293 cells in the context of virus infection (**Fig. 1a** and **Supplementary Fig. 1** online). Exogenous miR-93 and miR-124 were processed into pre-miRNA products and mature forms; this processing was unaffected by influenza A virus infection (**Fig. 1a**). Furthermore, endogenous miR-93 expression remained unchanged in both human and murine fibroblasts during the course of infection (**Fig. 1b**). Monitoring of miR-124–induced repression of a luciferase reporter containing known miR-124 target sequences[13] showed that miR-124 specifically suppressed 90% of luciferase activity; this inhibition was not disrupted by the presence of influenza A virus or NS1, the nonstructural RNA-binding protein responsible for disrupting many cellular processes (**Fig. 1c**,**d**)[14]. These data corroborate evidence[15] that influenza A virus and NS1 allow for proper miRNA biogenesis and gene silencing in mammalian cells and suggest that viral attenuation by MRE incorporation is feasible.

However, application of this strategy to influenza A virus is hindered by the fact that the viral mRNA terminates shortly downstream of the stop codon, and disruption of this region can result in packaging and replication defects[16]. We therefore incorporated MREs into the viral open reading frame. To design a virus that would be attenuated in humans and mice without reducing vaccine yield in eggs, we identified miRNA species that are not expressed in chicken but are ubiquitous in both murine and human lung tissue. Published deep sequencing results of miRNA profiles from *Gallus gallus*, *Mus musculus* and *Homo sapiens* identified miR-93 as an ideal candidate (**Supplementary Fig. 2a** online)[17–20]. The ubiquity of miR-93 in *Mus musculus* and *Homo sapiens* was corroborated by northern blot analysis, RT-PCR and RNA *in situ* data from *Mcm7*, which encodes miR-93 as part of an intron (**Supplementary Fig. 2b–d**).

To incorporate miR-93 target sites into influenza A virus, we identified regions in the viral genome that maintain high amino acid conservation between circulating strains but demonstrate plasticity at the RNA level. We chose the highly conserved segment five, which encodes nucleoprotein. We reasoned that this conservation would reduce the likelihood of escape mutants while allowing for DNA rescue of a diverse set of recombinant influenza A virus strains[2]. We identified two stretches of RNA that could be changed into miR-93 target sites without modifying the physical properties of the amino acids encoded at these sites (**Fig. 2a**). To ensure efficient and effective targeting, and to further decrease the possibility of revertants, we designed two near-perfect MREs at positions 225 (site 1, 93NP1) and 818 (site 2, 93NP2) of A/Puerto Rico/8/34 (A/PR/8/34), a laboratory H1N1 strain of influenza A virus that was adapted from a circulating 1934 strain[21] (**Fig. 2a**). Although codon usage at these sites is not conserved among circulating H1N1, H3N2 and H5N1 strains, these sites consistently maintain the same hierarchical class of amino acids,

[1]Microbiology Graduate School Training Program, [2]Department of Microbiology, [3]Global Health and Emerging Pathogens Institute, Mount Sinai School of Medicine, New York, New York, USA. Correspondence should be addressed to B.R.t. (benjamin.tenoever@mssm.edu).

Received 26 November 2008; accepted 27 April 2009; published online 31 May 2009; doi:10.1038/nbt.1542

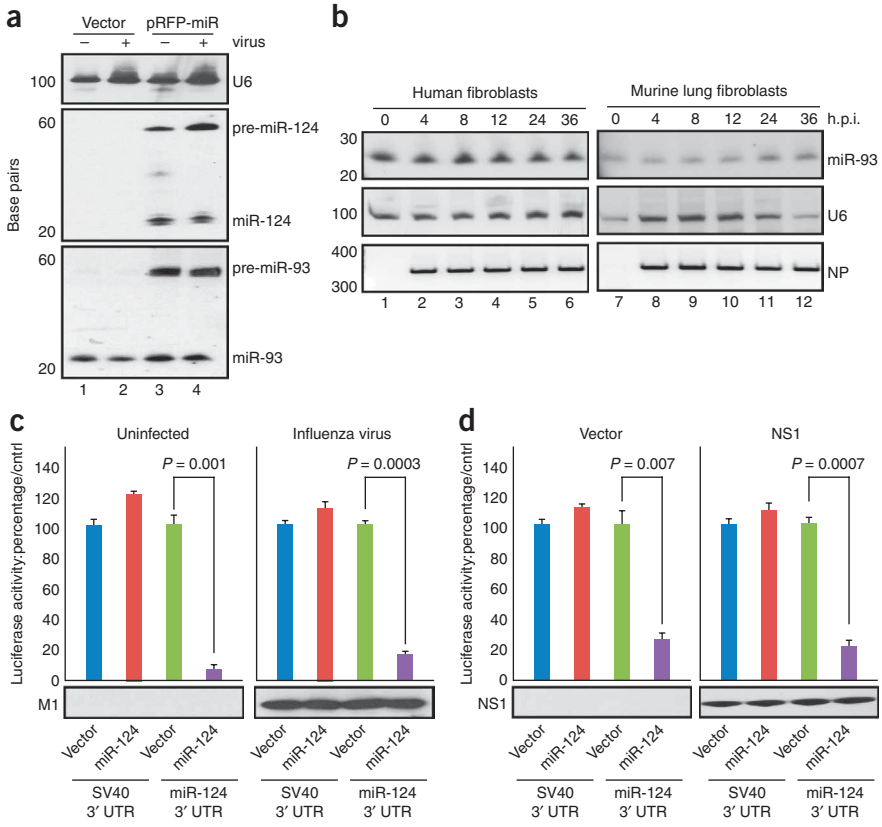


Figure 1 Influenza A virus infection does not disrupt cellular miRNA function. (**a**) Northern blot of exogenous miR-93, miR-124 and U6 small nuclear (sn)RNA after mock or influenza A virus infection (multiplicity of infection (MOI) = 10) of HEK293 cells. (**b**) Northern blot of endogenous miR-93 and U6 snRNA after influenza A virus infection (MOI = 3) for hours indicated in human and murine fibroblasts. NP, nucleoprotein. h.p.i., hours post infection. (**c**) HEK293 cells expressing either vector or miR-124 and luciferase reporter constructs containing control SV40 or miR-124 MRE-containing 3′ UTRs. Cells were infected with influenza A virus (MOI = 1) and subsequently measured for luciferase activity. (**d**) HEK293 cells expressing influenza A virus NS1, miR-124 and luciferase reporter constructs as described in **c**. For **c** and **d**, firefly luciferase activity was normalized to *Renilla* luciferase; data are the means of three independent experiments; error bars represent ± s.d. Western blots beneath each graph depict the expression of matrix 1 (M1) and NS1 proteins.

that is, the side chain at these positions maintains its hydrophobicity, polarity and/or charge (**Supplementary Fig. 3** online).

Modifying the coding region of segment five at sites one and two generated high-affinity miR-93 binding sites that exceed the complementarity of canonical MREs[8] (**Fig. 2a**). However, these modifications resulted in three amino acid substitutions, I63L, S262T and I265L, all of which remained within their hierarchical order and thus still resembled the natural variation observed at these loci.

As three amino acid substitutions were required to generate MRE-containing nucleoprotein, we designed a control virus, called parental (PRNTL), that contained these three amino acid substitutions and included additional mutations that disrupt miR-93 binding (**Fig. 2a**). PRNTL allowed us to distinguish phenotypic differences due to miRNA processing from indirect effects mediated by changes in nucleoprotein structure.

To ascertain whether nucleoprotein function was compromised by the I63L, S262T and I265L substitutions, we used an antisense reporter construct to determine polymerase function[22]. Exogenous expression of the three main polymerase components PB1, PB2, PA and either the wild-type or the PRNTL nucleoprotein was used to measure RNA-dependent RNA-polymerase (RdRp)-driven luciferase expression. The data showed that the three conservative substitutions made to nucleoprotein did not substantially affect its function, but did result in 20% reduction of RdRp activity (**Fig. 2b**). To determine the effects of these amino acid substitutions in the context of virus infection, we compared an H1N1 strain (A/PR/8/34) encoding PRNTL nucleoprotein to wild type (WT) A/PR/8/34. PRNTL and WT virus infection were similar in cell culture; however, *in vivo* infections indicated that PRTNL was attenuated (**Fig. 2c**). This attenuation may reflect the decreased RdRp activity of PRNTL, which would be more apparent under the selective pressure of an *in vivo* infection than in immortalized cell culture.

To determine how MRE-containing viruses compare to the PRNTL virus *in vivo*, we infected mice with increasing titers of either PRNTL or 93NP1/2 recombinants. Despite lethality in both cohorts at titers $>1 \times 10^5$ plaque forming units (PFU), mortality was restricted to infections with the PRNTL strain at intranasal inocula of 10^4 and 10^3 PFU, generating a greater than two-log difference in lethal dose between PRNTL and the MRE-containing 93NP1/2 (**Fig. 2d**); median lethal dose (LD_{50}) of 1.7×10^3 and 2.15×10^5 for PRTNL and 93NP1/2, respectively[23]).

Next, we determined whether MRE-containing viruses are attenuated *in ovo*. We inoculated 10-d-old embryonated chicken eggs with 100 PFU of wild-type A/PR/8/34/H1N1 (WT), PRNTL, 93NP1, 93NP2 or 93NP1/2 (**Fig. 2e**). Two days after infection, allantoic fluid was harvested and titers of $\sim 10^8$ PFU/ml were observed for each of the five strains, suggesting no attenuation *in ovo* (**Fig. 2e**). Furthermore, to illustrate the versatility of this technology, we used the MRE-containing nucleoprotein segments to rescue H5N1 6:2 reassortants, generating viruses antigenically recognized as A/Vietnam/1203/04/H5N1 via modified hemagglutinin and neuraminadase gene expression[24]. As with the H1N1 strains, the reassortant viruses demonstrated no attenuation in eggs, growing to titers of $\sim 10^8$ PFU/ml (**Fig. 2f**).

To determine whether species-specific attenuation was a result of miRNA-mediated gene silencing, we infected *Dicer*$^{-/-}$ murine fibroblasts, which do not process miRNAs (**Supplementary Fig. 2b**), and wild-type murine fibroblasts with A/PR/8/34-based PRNTL, 93NP1, 93NP2 or 93NP1/2 strains (**Fig. 3a** and **Supplementary Fig. 4** online). In wild-type fibroblasts, PRNTL virus produced abundant levels of hemagglutinin, 93NP1 and 93NP2 showed mild attenuation of hemagglutinin production, and 93NP1/2 produced no hemagglutinin (**Fig. 3a**). Attenuation was attributed to miRNA targeting as these same viral strains replicated to high, similar levels in the absence of *Dicer* (**Fig. 3a**). To ensure that this attenuation was miR-93 specific, we pretreated cells with Mercury locked nucleic acid (LNA) anti-miR-93 or scrambled RNA and subsequently infected cells with PRNTL, 93NP1 or 93NP1/2 (**Fig. 3b**). In the presence of anti-miR-93, 93NP1/2 levels were greater than twice that observed with scrambled

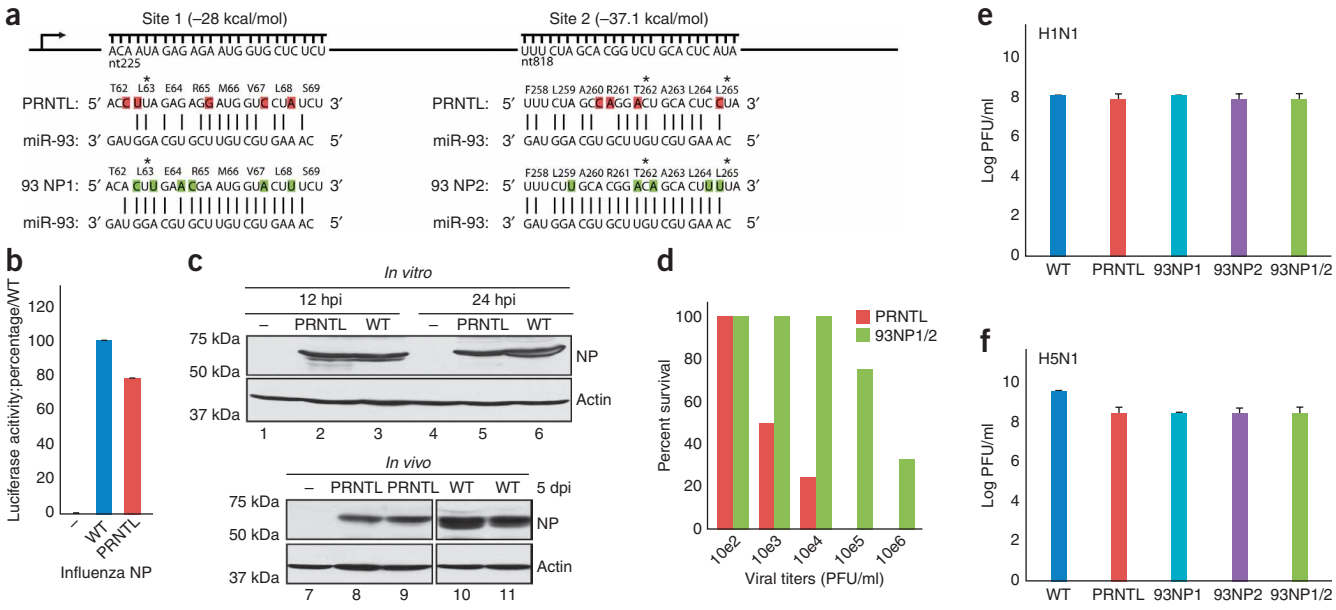


Figure 2 Species-specific attenuation of H1N1 and H5N1 influenza A viruses. (**a**) Schematic of RNA base substitutions generated to transform site 1 and site 2 in parental (PRNTL) or miR-93–targeted (93NP1/2) sites. *, denotes amino acid substitutions. Colored blocks denote nucleotide substitutions. Mean free energy is included as kcal/mol for each respective optimized site. (**b**) Influenza A virus polymerase-based firefly luciferase reporter assay in the context of no nucleoprotein (NP), wild-type (WT) NP or PRNTL NP. Data are the means of three independent experiments, normalized to *Renilla* luciferase; error bars represent ± s.d. (**c**) Western blot analysis of HEK293 cells or lung extract from BALB/c mice treated with WT influenza A/PR/8/34 or PRNTL viruses. Immunoblots of nucleoprotein and actin are shown. (**d**) Intranasal *in vivo* toxicity of PRNTL and 93NP1/2 viruses in BALB/c mice ($n = 4$/dose/strain). Inoculating titers given in PFU/ml. (**e**) Viral titers from PRNTL and MRE-containing H1N1 influenza A virus infections of 10-d-old embryonated chicken eggs. Titers were determined by hemagglutination and plaque assay from allantoic fluid and expressed as PFU/ml. Data are the means of four independent experiments. (**f**) Same as in e for H5N1 reassortant viruses.

control, indicating that attenuation of recombinant virus was a result of miR-93–specific suppression of nucleoprotein.

To further characterize the attenuation mechanism, we performed *in vitro* and *in vivo* kinetic, quantitative experiments comparing the levels of nucleoprotein and mRNA (**Fig. 3c**,**d**). As nucleoprotein is critical for copying template RNAs to form both mRNA and viral (v)RNA[25], we infected HEK293 cells and extracted total protein and RNA at 12 and 24 h after infection to determine whether decreases in nucleoprotein levels preceded the loss of mRNA. In cells infected with 93NP1/2, levels of nucleoprotein mRNA were elevated whereas protein levels were very low. This pattern of high nucleoprotein mRNA and low nucleoprotein was also observed *in vivo* 48 h after infection (**Supplementary Fig. 5** online). These results suggest that, despite the extensive complementarity, miR-93–mediated attenuation of influenza A virus was the result of translational repression. The overproduction of nucleoprotein mRNA may reflect the role of unbound nucleoprotein in the elongation of vRNA chains, loss of which biases the switch from transcription to replication[25].

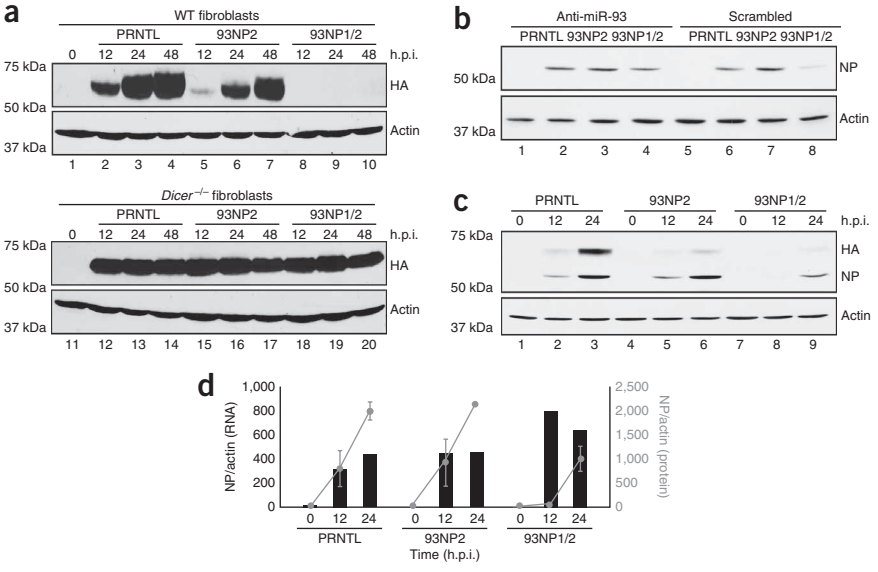


Figure 3 Characterization of miRNA-mediated attenuation of influenza A virus. (**a**) Western blot analysis of A/PR/8/34-based PRNTL, 93NP2 and 93NP1/2 infections harvested at hours indicated in wild-type (WT) and *Dicer*$^{-/-}$ murine fibroblasts. Immunoblots of hemagglutinin and actin protein levels are shown. h.p.i., hours post infection. (**b**) Western blot analysis of HEK293 cells treated with Mercury LNA anti-miR-93 or scrambled RNA oligonucleotides and infected as in **a**. Immunoblots of nucleoprotein and actin are depicted. (**c**) Western blot analysis of HEK293 cells infected as described in **a**. Immunoblots for hemagglutinin and nucleoprotein are shown. (**d**) Left-hand vertical axis: RNA levels determined by quantitative RT-PCR, standardized to actin and represented as copy number; error bars represent ± s.d. Right-hand vertical axis: quantification of protein levels by densitometry; nucleoprotein levels standardized to actin and represented in arbitrary units.

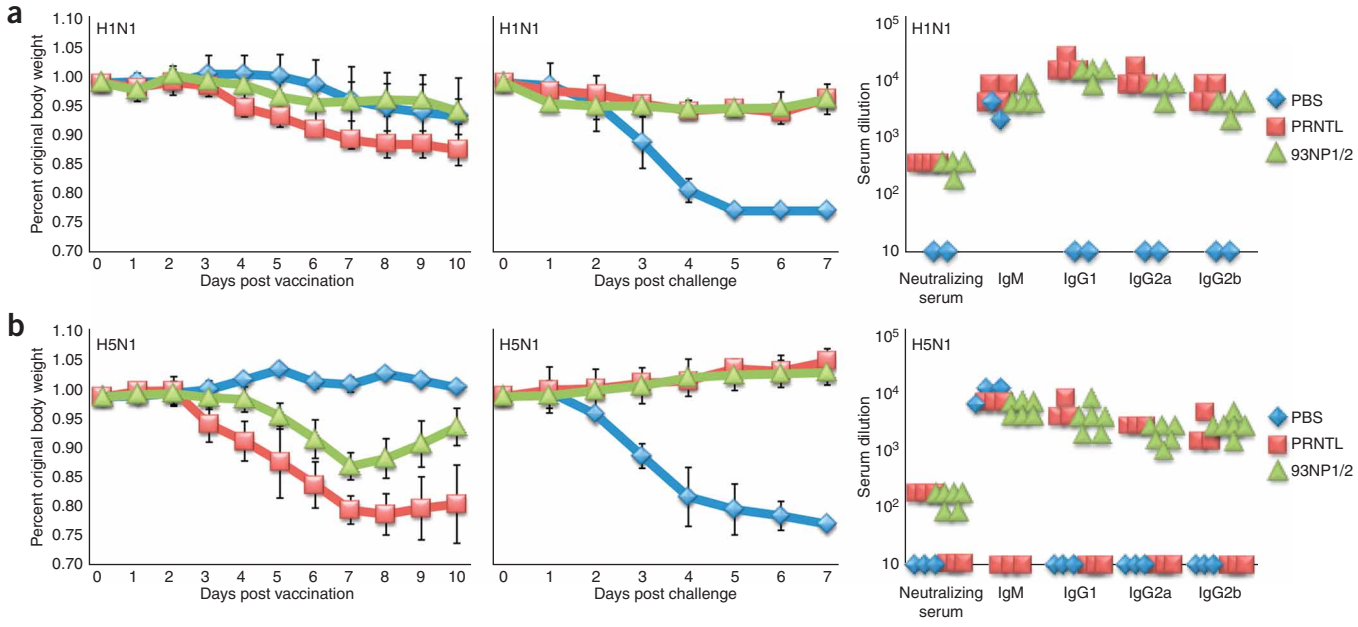


Figure 4 MRE-containing influenza A viruses as live attenuated vaccines. (**a**) Left: percent original body weight for mice vaccinated intranasally with 10e3 PFU of PRNTL or 93NP1/2 H1N1 viruses. Middle: percent original body weight after lethal challenge. Data represent the mean of each cohort ($n = 4$), errors bars are ± s.d. Right: antibody response after challenge, as determined by lowest serum dilution of positive hemagglutinin inhibition and immunoglobulin response as measured by enzyme-linked immunosorbent assay. (**b**) Same as in **a**, with H5N1 reassortant PRNTL or 93NP1/2 vaccinations.

Additional *in vivo* characterization of the MRE-containing influenza viral strains confirmed no changes in viral tropism and a normal cellular response to infection, demonstrating robust cytokine and transcriptional profiles (**Supplementary Fig. 6** online). To evaluate the possible generation of escape mutants, we analyzed RNA sequences from both *in vitro* and *in vivo* infections. Notably, we were unable to isolate escape mutants, suggesting that the flexible nature of miRNA targeting combined with the opposing rigid conservation of nucleoprotein may prevent reversion and add to the safety of this LAIV strategy.

To ascertain whether miR-93–seeded strains serve as effective LAIVs, we performed studies in mice with the A/PR/8/34 H1N1 PRNTL and 93NP1/2 recombinants (**Fig. 4a**). Inoculation of the PRNTL strain resulted in >10% weight loss compared with administration of 93NP1/2 or PBS. Furthermore, lethal challenge of these mice, 21 d after 93NP1/2 vaccination, resulted in 100% survival and induction of a robust repertoire of antibodies, including IgM, IgG1, IgG2a and IgG2b (**Fig. 4a**).

To further test the versatility of this MRE-based vaccine strategy, we inoculated mice intranasally with the miR-93–seeded H5N1 reassortants, which demonstrated no attenuation *in ovo* (**Fig. 2e**). Unlike the H1N1 vaccinations, H5N1 PRNTL resulted in 50% mortality and an average 20% loss in body weight (**Fig. 4b**). In contrast, all mice vaccinated with MRE-containing H5N1 survived, with only a small loss in body weight (**Fig. 4b**). Twenty-one days after vaccination, mice were challenged with a lethal dose of H5N1. Mock-vaccinated animals showed rapid weight loss and 100% mortality. In contrast, mice vaccinated with MRE-containing H5N1 displayed no signs of morbidity, indicating complete protection (**Fig. 4b**). Furthermore, serum samples from these mice were positive for neutralizing activity against wild-type H5N1 virus and, as in the case of H1N1 vaccinations, had high levels of IgM, IgG1, IgG2a and IgG2b (**Fig. 4b**).

In this report, we describe the effectiveness of harnessing the endogenous miRNA silencing machinery to achieve species-specific attenuation of influenza A virus. The MRE-containing LAIVs are attenuated in mammals while maintaining wild-type characteristics *in ovo*. Although the H1N1 reassortant MRE-containing vaccine described here does not protect against any influenza A strains currently in circulation, the H5N1 reassortant should be an effective vaccine for current H5N1 strains. This technology could be applied to any other influenza A strain. An advantage of MRE-based attenuation is that the degree of attenuation can be modulated by varying the number of MREs and/or the miRNA(s). This approach can also easily be adapted to tissue culture through the exploitation of miRNAs that are absent in select cell lines. Lastly, as the mechanism is distinct from temperature-based attenuation, this technology could be used in concert with FluMist as a means of increasing vaccine safety and extending the target demographic to include those presently excluded on the basis of their age.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturebiotechnology/.

Note: Supplementary information is available on the Nature Biotechnology website.

ACKNOWLEDGMENTS

This material is based upon work supported in part by the US Army Research Laboratory and the US Army Research Office under grant number 54677-LS-YIP. J.T.P. is supported by the NYU-MSSM Mechanisms of Virus-Host Interactions National Institutes of Health T32 training grant (no. AI007647-09). B.R.t. is supported in part by the Pew Charitable Funds. We thank A. Garcia-Sastre for giving us polyclonal A/PR/8/34 (MSSM, NY); G. Sen for ISG54 (Cleveland Clinic, OH, USA); A. Fernandez-Sesma (MSSM, NY) for primary human dendritic cell RNA; A. Tarakhovsky (Rockefeller University, NY) and D. O'Carroll (EMBL, Monterotondo, Italy) for *Dicer−/−* murine fibroblasts; E. Makeyev (Nanyang Technological University, Singapore) for firefly luciferase constructs containing miR-124 MREs and control SV40 3′ UTRs; and P. Palese (MSSM, NY) for pPol-I driven nucleoprotein vector, monoclonal nucleoprotein and A/PR/8/34 NS1. We thank members of the Palese and Garcia-Sastre labs for advice, reagents and comments during the course of this work.

AUTHOR CONTRIBUTIONS
J.T.P. and B.R.t. designed and wrote the manuscript. Animal studies were conducted by A.M.P. and M.A.C. Experiments were done by J.T.P., A.M.P. and M.H.L. Cloning and rescue of H5N1 reassortants was done by J.S.

Published online at http://www.nature.com/naturebiotechnology/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Stohr, K. The global agenda on influenza surveillance and control. *Vaccine* **21**, 1744–1748 (2003).
2. Neumann, G., Fujii, K., Kino, Y. & Kawaoka, Y. An improved reverse genetics system for influenza A virus generation and its implications for vaccine production. *Proc. Natl. Acad. Sci. USA* **102**, 16825–16829 (2005).
3. Ulmer, J.B., Valley, U. & Rappuoli, R. Vaccine manufacturing: challenges and solutions. *Nat. Biotechnol.* **24**, 1377–1383 (2006).
4. Brown, B.D., Venneri, M.A., Zingale, A., Sergi Sergi, L. & Naldini, L. Endogenous microRNA regulation suppresses transgene expression in hematopoietic lineages and enables stable gene transfer. *Nat. Med.* **12**, 585–591 (2006).
5. Barnes, D., Kunitomi, M., Vignuzzi, M., Saksela, K. & Andino, R. Harnessing endogenous miRNAs to control virus tissue tropism as a strategy for developing attenuated virus vaccines. *Cell Host Microbe* **4**, 239–248 (2008).
6. Kelly, E.J., Hadac, E.M., Greiner, S. & Russell, S.J. Engineering microRNA responsiveness to decrease virus pathogenicity. *Nat. Med.* **14**, 1278–1283 (2008).
7. Edge, R.E. *et al.* A let-7 MicroRNA-sensitive vesicular stomatitis virus demonstrates tumor-specific replication. *Mol. Ther.* **16**, 1437–1443 (2008).
8. Bartel, D.P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215–233 (2009).
9. Selbach, M. *et al.* Widespread changes in protein synthesis induced by microRNAs. *Nature* **455**, 58–63 (2008).
10. Baek, D. *et al.* The impact of microRNAs on protein output. *Nature* **455**, 64–71 (2008).
11. Bartel, D.P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297 (2004).
12. Artzi, S., Kiezun, A. & Shomron, N. miRNAminer: a tool for homologous microRNA gene search. *BMC Bioinformatics* **9**, 39 (2008).
13. Makeyev, E.V., Zhang, J., Carrasco, M.A. & Maniatis, T. The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. *Mol. Cell* **27**, 435–448 (2007).
14. Hale, B.G., Randall, R.E., Ortin, J. & Jackson, D. The multifunctional NS1 protein of influenza A viruses. *J. Gen. Virol.* **89**, 2359–2376 (2008).
15. Kok, K.H. & Jin, D.Y. Influenza A virus NS1 protein does not suppress RNA interference in mammalian cells. *J. Gen. Virol.* **87**, 2639–2644 (2006).
16. Palese, P. & Shaw, M. Orthomyxoviridae: the viruses and their replication. in *Fields Virology* vol. 2, edn. 5 (eds. Knipe, D.M *et al.*) 1648–1689, (Lippincott Williams and Wilkins, Philadelphia, 2006).
17. Burnside, J. *et al.* Deep sequencing of chicken microRNAs. *BMC Genomics* **9**, 185 (2008).
18. Williams, A.E., Perry, M.M., Moschos, S.A. & Lindsay, M.A. microRNA expression in the aging mouse lung. *BMC Genomics* **8**, 172 (2007).
19. Landgraf, P. *et al.* A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129**, 1401–1414 (2007).
20. Barad, O. *et al.* MicroRNA expression detected by oligonucleotide microarrays: system establishment and expression profiling in human tissues. *Genome Res.* **14**, 2486–2494 (2004).
21. Schickli, J.H. *et al.* Plasmid-only rescue of influenza A virus vaccine candidates. *Philos Trans. R. Soc. Lond. B Biol. Sci.* **356**, 1965–1973 (2001).
22. Hoffmann, H.H., Palese, P. & Shaw, M.L. Modulation of influenza virus replication by alteration of sodium ion transport and protein kinase C activity. *Antiviral Res.* **80**, 124–134 (2008).
23. Reed, L. & Muench, H. A simple method of estimating fifty percent end points. *Am. J. Hyg.* **27**, 493–497 (1938).
24. Li, S. *et al.* Priming with recombinant influenza virus followed by administration of recombinant vaccinia virus induces CD8+ T-cell-mediated protective immunity against malaria. *Proc. Natl. Acad. Sci. USA* **90**, 5214–5218 (1993).
25. Shapiro, G.I. & Krug, R.M. Influenza virus RNA replication in vitro: synthesis of viral template RNAs and virion RNAs in the absence of an added primer. *J. Virol.* **62**, 2285–2290 (1988).

ONLINE METHODS

Generation of recombinant virus. Human embryonic kidney cells (HEK293) were transfected with parental or mutant RNA-polymerase I–dependent nucleoprotein (pPol-I-NP) constructs, which produce vRNA transcripts, along with a Pol II–dependent nucleoprotein construct (pCAGGs-NP) and seven pDZ constructs that encode bidirectional vRNA and mRNA, as previously described[26–28]. Cells were harvested 24 h after transfection and injected into the allantoic fluid of 10-d-old embryonated chicken eggs. Live virus was isolated 48 h after infection and quantified by hemagglutinin assay using chicken red blood cells in Alsevers (CBT Farms) and plaque assay in Madin-Darby canine kidney (MDCK) cells. Plaque assays were limited to parental and wild-type virus as MDCK cells express significant levels of miR-93 (**Supplementary Fig. 2b**). H5N1 recombinant influenza A viruses were generated using internal components of H1N1 (A/PR/8/34) and the hemagglutinin and neuraminadase segments of H5N1 (A/Vietnam/1203/04). The hemagglutinin segment was modified by the removal of the polybasic cleavage site, as previously described[28].

Infections in mice. Animal infections were performed in 5-week-old BALB/c mice (Taconic). Mice were put under general anesthetic and treated intranasally. Pathogenic studies were performed on cohorts of 3–4 mice/inoculating dose. Mice were weighed daily and euthanized if they lost >25% of original body mass. Vaccinations were performed using 1 × 10e3 PFU of MRE-containing H5N1 or H1N1. Three weeks after infection, mice were rechallenged intranasally with either 10 LD_{50} of H5N1 6:2 recombinant or 100 LD_{50} of H1N1 (A/PR/8/34). Mice were monitored daily for signs of morbidity and mortality.

ELISA and HI assay. Serum was obtained post mortem through intraocular bleeding and treated with cholera filtrate (Sigma) overnight at 37 °C. ELISA and hemagglutinin assays were performed as previously described[29].

miRNA expression and targeted luciferase vectors. The red fluorescent protein minigene (pRFP) expressing miR-124 has been described elsewhere[13]. For generation of pRFP-miR-93, a 500-bp genomic fragment containing the pri-miR-93 locus was isolated from mouse genomic DNA by PCR amplification with High Fidelity PCR Master kit (Roche) per the provided protocol, using forward 5′-TAGTGGTCCTCTCTGTGCTACCG-3′ and reverse 5′-ATTGAACA AAAATGGGGACTCCT-3′ primers. The resulting PCR product was subcloned into PCR2.1-TOPO (Invitrogen) according to the manufacturer's instructions, and subsequently cloned into the pRFP minigene via PmeI-SpeI sites. Firefly luciferase constructs containing miR-124 MREs and control SV40 3′ UTRs were a kind gift from E. Makeyev (Nanyang Technological University, Singapore)[13].

Passaging of virus and sequencing of nucleoprotein. Human lung epithelial cells (A549) were infected with PRNTL or MRE-containing H1N1 (MOI = 0.01) in the presence of TPCK trypsin. Cells were washed 2 h after infection in PBS and then replaced with complete media. One hundred microliters of replacement media was removed from the infected plates 15 min after changing the media and tested by plaque assay to ensure absence of virus from the original inoculum, and 24 h after infection, 10 µl of supernatant (from a total volume of 1.5 ml) was serially transferred to naive cells for a total of 10 passages. Ten days after infection, RT-PCR was performed on total RNA and nucleoprotein PCR products were cloned and sequenced. For *in vivo* studies, 5-week-old BALB/c mice were treated with virus as above. Five days after infection, total RNA was harvested, and used to clone nucleoprotein for sequencing. Depicted sequences represent >25 individual colonies per cohort.

Tissue culture and *in vitro* infections. HEK293, A549, murine lung fibroblasts and MDCK cells were grown in DMEM (Mediatech), supplemented with 10% FBS (JM Bioscience) and 1% penicillin/streptomycin (P/S, Mediatech), unless otherwise indicated. Human astrocytoma U373 cells were supplemented with an additional 10 mM HEPES (GIBCO). Jurkat cells were grown in RPMI (Mediatech), supplemented with 10% FBS, and 1% P/S. Primary human dendritic cell RNA was kindly provided to us by A. Fernandez-Sesma (MSSM, NYC). *Dicer*−/− murine fibroblasts were a kind gift from A. Tarakhovsky (Rockefeller University, NYC) and Donal O'Carroll (EMBL, Monterotondo), and were grown in DMEM supplemented with 15% FBS, 1% nonessential amino acids (GIBCO), and 1% P/S. *In vitro* infections were performed in serum-free medium for 1 h; inoculum was washed and replaced with fresh complete medium without trypsin for indicated times. pRFP expressing HEK293 cells were infected with WT influenza A/PR/8/34 (MOI = 10) for 24 h, and subsequently harvested for total RNA. Human and murine lung fibroblasts were infected with WT A/PR/8/34 (MOI = 3) and harvested at the times indicated. Murine wild-type and *Dicer*−/− fibroblasts were infected with PRNTL or MRE-containing A/PR/8/34 viruses (MOI = 1). For anti-miR-93 experiments, HEK293 cells were transfected with 20 nM of Mercury LNA oligonucleotides (Exiqon) specific for human miR-93 or scrambled control for 24 h preceding infection. Subsequently, cells were infected with PRNTL or MRE-containing A/PR/8/34 viruses (MOI = 0.1) and harvested 24 h post infection. Comparison of nucleoprotein and RNA levels after infection was performed in HEK293 cells (MOI = 0.1).

Quantitative-PCR and western blot. qPCR and analysis were performed by MSSM Quantitative Genomics core facility. RT-PCR and immunoblots were performed as recently described[29]. RT-PCR primers are listed below. Actin (Abcam), polyclonal A/PR/8/34 (a kind gift from A. Garcia-Sastre), monoclonal nucleoprotein (a kind gift from P. Palese, MSSM, NYC), IRF1 (Santa cruz, sc-640), STAT1 (Santa cruz, sc-417), and ISG54 (a kind gift from G. Sen) antibodies were all used at a concentration of 1 µg/ml and incubated overnight at 4 °C. Secondary mouse and rabbit antibodies (GE Healthcare) were used at a 1:5,000 dilution for 1 h at 25 °C. Full-length western blots are presented in **Supplementary Figure 7** online. miRNA RT-PCR primers include: Let7a: GTCCTGGCGCGGTGCTCT & TCTCTTGCTCCTTCCCT TGC; miR155: CATTTCAAGAACAACCTACCAGAGA & AAGTTTATCCAGC AGGGTGACTC; mIR16: TCTGATGTGAACACAAGGACATTCA & TTTCCAC CATCTTTACCCTGTTT3; miR93:GAAGCTCATGAGGCGTTACATAG & ATT GACCTGCCAGACATTGAG; miR181: CAACGGTTTCTGTCAGGATGAAT & AGGGGAACTGTGGTCACTATCAC; mIR21: TGCTTGGGAGGAAAATAAAC AAT & GACTCTAAGTGCCACCAGACAGA; tubulin: GCCTGGACCACAAGT TTGAC & TGAAATTCTGGGAGCATGAC.

pRFP and luciferase reporter transfections. Transfections for fluorescence confirmation of pRFP constructs were performed in HEK293s using 4 µg pRFP vector and Lipofectamine 2000 (Invitrogen), according to the manufacturer's instructions. Fluorescence was imaged 24 h after transfection. For subsequent infection with WT influenza A/PR/8/34, HEK293s were transfected as above, with 100 ng firefly luciferase 3′ UTR constructs, 10 ng constitutive *Renilla* luciferase, and 700 ng of miRNA pRFP construct and harvested 18 h after transfection and 12 h after infection (MOI = 1) and analyzed per the manufacturer's instructions (Dual-Luciferase Reporter Assay, Promega). For NS1 studies, HEK293 cells were cotransfected with pBluescript SK+ (Stratagene) or A/PR/8/34 NS1 (a kind gift from P. Palese, MSSM, NYC) in addition to 50 ng firefly luciferase 3′ UTR constructs, 10 ng constitutive *Renilla* luciferase, 350 ng miRNA pRFP constructs. Cells were harvested 24 h after transfection and analyzed as described above. RNA-dependent RNA polymerase activity of mutant nucleoprotein was accessed in HEK293 cells transfected with 250 ng of pDZ-NP-PRNTL1/2 or wild-type pDZ-NP and 100 ng pPol-I firefly luciferase and analyzed as previously described[22]. All firefly luciferase readings are expressed as a ratio to *Renilla* luciferase expression per sample, and averaged over three independent replicates.

Statistical analyses. Statistical analysis was performed using a two-tailed Student's *t*-test with an $n = 3–8$. $P < 0.05$ were considered significant, and error bars reflect ± s.d.

miRNA northern blot analysis. Northern blots in **Figure 1** were generated from total RNA and separated by PAGE (PAGE) with a 15% denaturing polyacrylamide gel containing 7.5 M urea and 1× TBE[14]. RNA was transferred to Hybond N+ membrane (Amersham) in 0.5× TBE, cross-linked and blocked overnight. Probes include: anti-miR-124: 5′-TGGCATTCACCGCGTGCCT TAA-3′, anti-miR-93: 5′-CTACCTGCACGAACAGCACTTTG-3′, and anti-U6: 5′-GCCATGCTAATCTTCTCTGTATC-3′. Probes were labeled using T4 polynucleotide kinase (Invitrogen) and [γ-^{32}-P]ATP (Perkin Elmer) and purified by Sephadex G-25 columns (GE Healthcare). Northern blots depicted in **Figure 2** were performed as previously described[30]. All northern blots portray a representative result that has been produced a minimum of three times.

Incorporation of MREs into influenza A/PR/8/34 NP. Sites within influenza A/PR/8/34 nucleoprotein with partial complementarity to miR-93 were identified using Bibiserv's RNAhybrid algorithm. Nearly full complementarity was achieved with 3–5 steps of site-directed mutagenesis using the QuickChange kit and protocol (Stratagene) on the pPol-I driven nucleoprotein vector (a kind gift from P. Palese, MSSM, NYC).

26. Quinlivan, M. *et al.* Attenuation of equine influenza viruses through truncations of the NS1 protein. *J. Virol.* **79**, 8431–8439 (2005).
27. Fodor, E. *et al.* Rescue of influenza A virus from recombinant DNA. *J. Virol.* **73**, 9679–9682 (1999).
28. Park, M.S., Steel, J., Garcia-Sastre, A., Swayne, D. & Palese, P. Engineered viral vaccine constructs with dual specificity: avian influenza and Newcastle disease. *Proc. Natl. Acad. Sci. USA* **103**, 8203–8208 (2006).
29. tenOever, B.R. *et al.* Multiple functions of the IKK-related kinase IKKepsilon in interferon-mediated antiviral immunity. *Science* **315**, 1274–1278 (2007).
30. Pall, G.S. & Hamilton, A.J. Improved northern blot method for enhanced detection of small RNA. *Nat. Protocols* **3**, 1077–1084 (2008).

doi:10.1038/nbt.1542

Career interrupted

Mari Paul

A sound game plan can help you roll with the punches in uncertain times.

"[A]ll knowledge is vain, save when there is work…" — Khalil Gibran

With the recent downturn in the economy and our business environment, many careers are going through a rough patch—not fun but ultimately survivable. How well we come out at the other end of this period can be helped immensely by alert management in just a few areas—namely our finances, skills and perspective. I called on Chanie Schwartz, a wealth manager for several pharma and biotech executives, to help me create a game plan for staying up or getting back on your feet in the ring.

Round one: a proactive career plan

First, realistically anticipate your future. Is the status of your group, and your place in your group, growing or declining in the company? Look for opportunities and put in the hours to take on new skills and responsibilities to grow both your value to the company and your résumé.

Second, revisit your dreams and passion. Determine what your next growth step is, using a career counselor if you can, and develop your action plan to achieve it. If possible, map out in advance your first full-time week of career creation as well as your personal goals for an anticipated six months of downtime. Now is the time to start writing the résumé you will need.

Finally, work on your networking and visibility, focusing in particular on your next career step. Take all conference, speaking and writing invitations possible and attend as many seminars as your employer will sanction. There are a whole group of additional supporters of your future you have yet to meet.

Mari Paul is the founder of Life Science Leaders, San Francisco, California, USA.
e-mail: mari@lifescienceleaders.com

Round two: a proactive financial plan

Mari Paul: The pharma and biotech industries are a long way from the mortgage crisis. How severely do you think we are being affected?

Chanie Schwartz: Plenty. A surprising number of high-income earners live paycheck to paycheck. For executives like this who are financially stretched or overcommitted, it's especially important to act now—well in advance of an unexpected event.

MP: How do you recommend that we prepare financially while on the job?

CS: First, focus on developing a comfortable relationship with your money. While you are still employed, can you live on just your base salary? Many people depend on bonuses and exercising stock options to meet their total annual cash demands. This may be the year that bonus structures change and also not the best time to sell stock.

Second, I suggest an 'emergency fund' of six to twelve months to cover the time between jobs. And third, many of you may be offered early retirement or decide to consult, and will want to have your retirement provided for and/or funds to start your new business.

MP: What's next?

CS: Then we start where we always do, with a comprehensive financial plan that incorporates information about income, employer stock holdings and other investments, and financial obligations—all in the context of your current requirements and your long-term financial and personal goals. This process often includes closely reviewing benefits and cash flow, evaluating an expected severance package and determining whether early retirement may be an option.

MP: This is more work than the usual advice we're used to hearing, which is "save, save, save." This is an information-gathering challenge, as well as a lot of planning, especially since many of us are putting in extra hours on the job as it is.

CS: Absolutely. You need more help than just tax advice. You first need a wealth manager to help develop the long-term strategy that will help maximize your financial security. Wealth managers bring together a team of professionals who will look after all of our complicated, interrelated financial issues. This team includes a CPA, along with an estate planning attorney, insurance agent, personal banker and others. This network of highly competent professionals generates quick, accurate analyses so that you can move ahead quickly with important decisions. And should you be terminated, hire an employment lawyer to look over your severance contract and negotiate your severance agreement.

MP: Is there anything a wealth manager would say about attitude?

CS: According to a psychiatrist I work with who focuses on executive stress, the best thing to do is stay active and tell everyone you know that you are looking for a job. That will help your survival skills.

Round three: the winner's edge

A great career that you can be passionate about is worth every effort. It's about a great journey more than the goal. Passion places money secondary to satisfaction, and Chanie and I agree that this game is about a lot more than money. The real win is the many rewards you get from using your natural gifts.

You can do some simple things that remind you of your self worth, such as crafting a personal statement of how you believe in yourself or making it a point to surround yourself with people who believe in you. And keep your winner's edge by reminding yourself that you have many talents that a business needs, and of the tremendous value that your passion, experience and knowledge provide to your industry.

PEOPLE

EnviroGene (Tredomen, UK) has named **Mark Chadwick** (left) as CEO. He joins the company from chemistry services provider Excelsyn and was previously involved in commercializing a novel drug discovery technology at Cambridge Genetics, which was later acquired by medicinal chemistry services provider BioFocus.

Alex Korda, chairman of EnviroGene, comments: "After two years of technology development, EnviroGene is preparing to move beyond field trials and into commercialization. We are delighted to welcome Mark as our CEO, and believe that his strong technical background and commercial acumen, along with entrepreneurial startup and small company experience, are ideally suited to driving the next phase of growth at EnviroGene."

Privately held Archemix (Cambridge, MA, USA) has announced the appointment of **Kenneth M. Bate** as president and CEO. Bate has more than two decades of leadership experience in the industry, most recently as president and CEO of NitroMed. He also held the positions of executive vice president, head of commercial operations and chief financial officer (CFO) of Millennium Pharmaceuticals and as vice president of sales and marketing and CFO at Biogen (now Biogen Idec).

SuperGen (Dublin, CA, USA) has announced the resignation of chief medical officer **Gregory Berk**. **Michael D. Young**, a current member of the SuperGen board of directors and chairman and CSO of Strategic Healthcare Development, will serve as SuperGen's interim chief medical advisor until a permanent replacement for Berk is appointed.

Avila Therapeutics (Waltham, MA, USA) has announced that **Katrine S. Bosley** has been appointed as CEO. Bosley joins the company from Adnexus, where she served as vice president, business development and later as vice president, strategic operations.

Patrick Fabbio (left) has joined Ikano Therapeutics (Saddle Brook, NJ, USA) as CFO. He previously held senior financial, transactional and operational positions with Sanofi-Aventis, where he most recently served as senior director, US life cycle management base business.

Privately held veterinary bioscience company Imulan BioTherapeutics (St. Joseph, MO, USA) has named **Daniel Gingerich** as CSO. Gingerich previously held senior research positions with a variety of human and veterinary pharmaceutical and life science companies including Stolle Milk Biologics, Fort Dodge Laboratories and Bristol Myers Animal Health.

Corey Goodman has resigned from his position as president of Pfizer's Biotherapeutics and Bioinnovation Center (S. San Francisco, CA, USA) just 19 months after his hiring as part of Pfizer's effort to expand its push into biotech products. Goodman co-founded biotech companies Exelixis and Renovis. In April, Pfizer announced the creation of a biotech division that would include Goodman's unit and the biotech research arm of Wyeth, which Pfizer is acquiring for $63 billion. **Mikael Dolsten**, now president of Wyeth research, was named to lead the new division. Goodman had previously reported directly to Pfizer CEO Jeffrey Kindler.

Ann Hayes has been appointed as a member and chair of Celentyx's (Birmingham, UK) board of directors. She spent 22 years at GlaxoWellcome, leaving in 2001. Since then, she has co-founded three companies—Ionix Pharmaceuticals, Therasci and Theradeas.

Daniel J. Lerner has been appointed chief medical officer of TYRX (Monmouth Junction, NJ, USA). He brings over 18 years of industry and clinical experience to TYRX, most recently as senior vice president for Foxhollow Technologies.

Joel Martin has been appointed president and CEO of privately held Altair Therapeutics (San Diego). Martin was a member of the early scientific team at Isis Pharmaceuticals, which pioneered antisense technology and discovered AIR645, which it licensed to Altair. Most recently, he was a partner at venture capital firm Forward Ventures.

ChemGenex Pharmaceuticals (Melbourne, Australia and Menlo Park, CA, USA) has announced the retirement of **Rick Merrigan**, the company's CFO and company secretary. He is to be succeeded in both positions by **James Campbell**, who has more than 20 years of experience in scientific research, research management, management consulting and venture capital, most recently as ChemGenex's COO.

Vertex Pharmaceuticals (Cambridge, MA, USA) has promoted **Peter Mueller**, executive vice president, drug innovation and realization and CSO, to executive vice president, global research and development and CSO. Mueller's responsibilities will cover all of Vertex's global R&D, including clinical and nonclinical development, clinical operations, medical and regulatory affairs.

Molecular Insight Pharmaceuticals (Cambridge, MA, USA) has named **Daniel L. Peters** president, CEO and a member of the board of directors. He succeeds company founder **John W. Babich**, who will now serve as executive vice president, CSO and president of R&D and continue as a member of the board of directors. Peters most recently served as president and CEO of GE Healthcare's medical diagnostics business.

SkyePharma (London) has appointed **Thomas Werner** as a new nonexecutive director. Werner is a senior-level pharmaceutical executive with over 26 years experience, most recently as managing director and senior vice president of GlaxoSmithKline Germany. Before that, he held senior positions at GlaxoWellcome Germany and Bristol-Myers Squibb Germany. In addition, **Argeris (Jerry) Karabelas** will stand down from the SkyePharma board after eight years of service.

Novocell (San Diego) has announced the appointment of **John West** as president, CEO and a member of its board of directors. West was formerly CEO of Solexa from 2004 until its acquisition by Illumina in 2007.