VOLUME 28 NUMBER 4 APRIL 2010
www.nature.com/naturebiotechnology

nature biotechnology

THE SCIENCE AND BUSINESS OF BIOTECHNOLOGY

Antagomirs tackle metastatic breast cancer

Broader resistance to plant pathogens

Image analysis software adds an extra dimension

VOLUME 28 NUMBER 4 APRIL 2010

nature biotechnology

Light micrograph of cultured human breast cancer cells. Weinberg and colleagues show that targeting a pro-metastatic microRNA with antagomirs prevents the formation of new metastases in a mouse model of breast cancer (p 341). Credit: Cecil H. Fox/Science Photo Library

Isis' cholesterol-lowering antisense drug meets endpoints, p 295

EDITORIAL

NEWS

BIOENTREPRENEUR

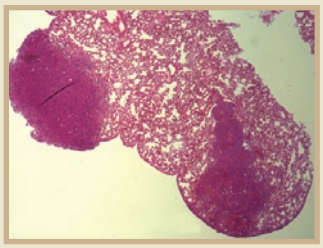
BUILDING A BUSINESS

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$3,520 (institution), US$4,050 (corporate institution). Canada add 5% GST #104911595RT001; Euro-zone: €202 (personal), €2,795 (institution), €3,488 (corporate institution); Rest of world (excluding China, Japan, Korea): £130 (personal), £1,806 (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

nature publishing group

Resistance to multiple bacterial pests, p 330



Targeting miRNAs curbs metastasis, p 341

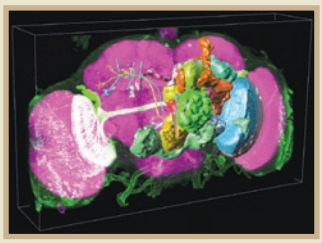


Image analysis in 3D, p 348

OPINION AND COMMENT

NEWS AND VIEWS

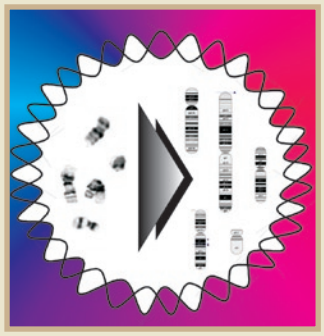
COMPUTATIONAL BIOLOGY

PRIMER

RESEARCH

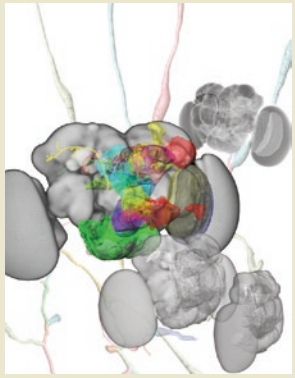
ARTICLES

LETTERS



Genetic profiling of hESCs, p 371

Now in 3D



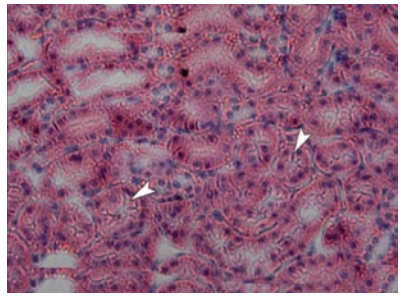
Microscopy images of biological structures, such as the fruitfly brain, typically capture many two-dimensional 'slices' of an object, but a number of studies would benefit from the ability to manipulate in real time a three-dimensional (3D) representation of these structures. Peng *et al.* describe the V3D software for accomplishing this task. Unlike existing software, which cannot handle the gigabyte-sized image stacks generated by today's microscopes, V3D makes it easier to view, rotate and enlarge a 3D representation that is generated on the fly from stacks of two-dimensional images. V3D also provides novel features for a user to directly select points in 3D space, which enables quantitative measurement and analysis. The basic functionality of V3D can be extended with specialized analysis algorithms or customized user interfaces. Peng *et al.* describe one such application for the neurosciences that allows them to accurately trace 100 neurites in images of the fruitfly brain and to begin constructing a 3D brain atlas. [**Articles, p. 348**] *CM*

Plug-and-play plant disease resistance

Stephane Genin

Disease resistance has previously been engineered in crops by expressing a receptor that activates defense responses upon recognizing a molecule specific to a particular pathogen. However, this resistance usually breaks down as part of the constant arms race between phytopathogens and plants. Zipfel and colleagues now demonstrate the potential of a possibly more effective weapon to shift the balance of power in this conflict. They express a so-called pattern-recognition receptor—a class of cell-surface molecules named because of their ability to recognize molecules conserved among several microbes—from *Arabidopsis thaliana* in both tomato and a relative of tobacco. Remarkably, this confers resistance of both species to a range of bacteria spanning four bacterial genera, presumably by activating defense signaling pathways not normally triggered in these species when they encounter these bacteria. As an added bonus, because the molecules recognized by most pattern-recognition receptors are almost always more essential for pathogen viability than effectors of 'gene-for-gene' resistance, this strategy may confer more durable resistance in field-grown crops. If the approach can be applied more generally, it could help to prevent massive crop losses while avoiding the environmental, health and financial costs associated with using pesticides. [**Letters, p. 365; News and Views, p. 330**] *PH*

Synthetic biology tunes drug delivery



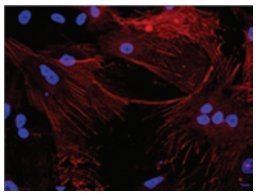
Genetic circuits have been engineered to exhibit different dynamic behaviors, but few have been put to practical use. Kemmer *et al.* take a step toward this goal by engineering a circuit that can sense, and then automatically regulate, the concentration of a metabolite in the bloodstream of a mammal. Their circuit consists of three components that are stably integrated in a transgenic cell line—a 'sensor' transcriptional repressor protein that permits expression of a therapeutic enzyme only when a metabolite in the bloodstream (uric acid) is overabundant, a 'responder' enzyme that converts uric acid into a form that can be more readily cleared from the blood, and a 'sensitizer' membrane transporter that allows sufficient amounts of uric acid to enter the cell. When implanted in a mouse model unable to process uric acid, the transgenic cells expressing this circuit maintain uric acid levels in the bloodstream within physiologically normal ranges and alleviate pathologic symptoms associated with diseases in which uric acid homeostasis is disrupted. [**Letters, p. 355**] *CM*

Antagomirs target cancer metastasis

Whereas primary tumors can be controlled using current therapies in many instances, effective treatments preventing the metastatic spread of cancer have proven more elusive. Weinberg and colleagues show that targeting a pro-metastatic microRNA in breast cancer can successfully reduce the formation of metastases in a mouse model of the disease. The authors use antagomirs—chemically modified antisense RNAs—to inhibit the function of miR-10b, a previously characterized pro-metastatic microRNA. After systemic application, ~80% fewer lung metastases are found in the antagomir-treated animals compared with mismatch control–treated animals. The antagomir has no effect on either the growth of the primary tumor or on already established metastases. Although miR-10b is expressed in many normal tissues, no toxicity is observed after treatment with the antagomir. [**Articles, p. 341; News and Views, p. 331**] *ME*

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

Growing hESCs in suspension



The development of cell therapies based on human embryonic stem cells (hESCs) will require new culture systems that allow these cells to be produced in large quantities. Existing culture methods are not well suited to this task because hESCs are generally grown as adherent monolayers, whereas suspension culture, which is needed for growth in bioreactors, tends to induce differentiation. Reubinoff and colleagues now show that hESCs can be propagated in suspension using a culture medium that combines multiple components of existing formulae. They also demonstrate the derivation of three new lines from floating inner cell masses, an achievement that promises to simplify and standardize the protocols for generating hESCs. [**Letters, p. 361**] *KA*

Patent roundup

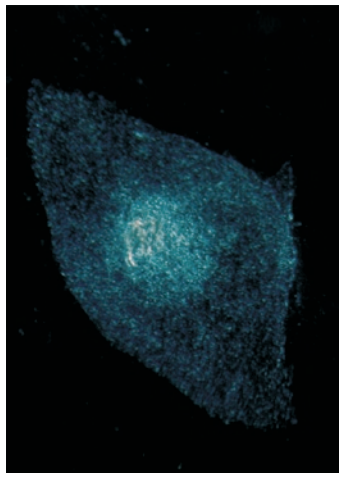
The US Patent and Trademark Office (Washington, DC) has issued a patent for detection of RNA-mediated gene silencing to Sir David Baulcombe of the University of Cambridge, and Andrew Hamilton, University of Glasgow, with implications for plant and mammalian research. **[News in Brief, p. 300]** *LM*

A Human Genome Sciences (Rockville, MD) patent has been struck down by a UK Court of Appeal for lack of industrial application—the first British case to deal with the validity of a gene sequence patent. **[News in Brief, p. 300]** *LM*

US patent reform legislation is once again on the horizon, aided by the specter of excessive litigation raising costs for industry. Mills and Tereskerz study stem cell patents to determine the rate of patent litigation and describe the litigation occurring in this new core technology. **[Patent Article, p. 325]** *MF*

Recent patent applications in high-throughput drug screening. **[New patents, p. 329]** *MF*

Finer map of hESC chromosomal changes



Human embryonic stem cells (hESCs) are known to acquire chromosomal abnormalities during long-term culture, a finding that raises safety concerns for cells to be used in patients. Lahesmaa and colleagues have carried out the highest-resolution genomic analysis to date of hESCs. Seventeen lines are analyzed using an Affymetrix (Santa Clara, CA, USA) array containing >900,000 probes for single nucleotide polymorphisms and >900,000 probes for copy number variations. The authors identify >800 copy number variations as well as sites of loss of heterozygosity, which has not been observed previously in hESCs. They also use transcriptional profiling to identify genes whose expression is altered as a result of culture-induced chromosomal changes. [**Resource, p. 371**] *KA*

Next month in nature biotechnology

- Qualification of biomarkers for kidney toxicity
- Insect epigenetics
- *De novo* transcriptome assembly
- Human ESC culture on defined surfaces

nature biotechnology

www.nature.com/naturebiotechnology

nature biotechnology

What health reform means for innovation

Healthcare reform will not only boost biotech investment by massively expanding the US drug market, but also change the dynamics of biotech innovation in the longer term.

Last month, President Obama kept his election promise and signed into law the most comprehensive piece of legislation concerning healthcare provision in the United States since the introduction of Medicare by Lyndon Johnson in 1965. Positioned as pro-industry and pro-recovery, the reform had become politically viable, despite equivocal support from large parts of the US Congress and the American people. In return for supporting the bill and stumping up $90 billion in fees and discounts on Medicaid and Medicare pricing, the drug industry receives tax breaks, a biosimilars pathway and a massively expanded drug market. In the longer term, the legislation promises to radically transform the economic drivers for biotech innovation.

The legislation—officially, the Patient Protection & Affordable Care Act (H.R. 3590) and the Health Care & Education Affordability Reconciliation Act of 2010 (H.R. 4872)—is less about healthcare reform than health insurance reform. It contains a scattering of measures that immediately benefit life science companies. The Biotechnology Industry Organization (BIO) has made much of the adoption of a 12-year exclusivity term for biosimilars, which for reasons of political expediency, was left intact despite the preference of the Obama Administration and generics manufacturers for a shorter brand monopoly. Referring to the Therapeutic Discovery Project Credit—a measure that gives tax credits (or grants for non-tax-paying companies) equivalent to half the cost of investments in "qualified therapeutic discovery projects" undertaken in fiscal 2009 or 2010—BIO also trumpeted the fact that the legislation will "provide some financial relief to research-intensive, small biotechnology companies." What qualifies is vague, but potentially the credit can apply to any preclinical, clinical, diagnostic or technology development undertaken by companies with <250 employees. The program is worth $1 billion ($500 million a year)—a large sum, although still only 11% of the annual R&D spending of public biotech companies with <250 employees.

Perhaps the biggest immediate boon to biotech, though, lies in the expansion of the market for prescription drugs. Having an additional 32 million people enter the world's largest healthcare market will provide an immediate stimulus to health-related businesses. Industry newsletter *The RPM Report* estimates that the new legislation could result in $115 billion in new business over 10 years. Given that drug companies dodged the bullets of drug price controls and drug importation, this means that extraordinarily high profits will likely continue for the sector, at least for a while, and investment will continue to be attracted to drug-related innovative biotech.

In the longer term, however, the most important aspect of the reform package lies in the enfranchisement of the previously uninsured and progress towards universality of healthcare access.

A new insurance pool will mean that people with pre-existing conditions will be able to obtain affordable health insurance. Furthermore, insurers will not be allowed to deny coverage to children because of pre-existing illness, and they won't be able to drop insurance coverage when a person becomes ill. Provisions in insurance plans that limit 'lifetime' coverage to some arbitrary age or impose annual limits on benefits will become illegal. In short, US insurers now have to insure and support the sick and vulnerable. Consequently, they need to assimilate the potential financial burden within their business models.

Until now, US health insurance profitability has been about circumventing the difficult and expensive burden of dealing with sick people. Insurers have treated illness as a type of 'insurance excess': they either capped their outlays or shifted sick people outside the insurance system. Now that this has been disallowed, insurers will have to find other ways of dealing with people with disease. At face value, at least, the legislation requires them to embrace the sick wholeheartedly and for as long as they are sick. And the pressure on insurers to reduce the costs of these patients has huge implications for life science innovation.

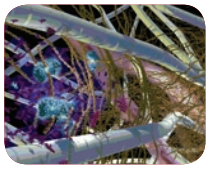
The recently passed legislation does relatively little to reduce the cost of biotech treatments. Although it does create a pathway for approval of biosimilar products, the prolonged market exclusivity period and burden of clinical proof make it unlikely that the biosimilars market will be sufficiently attractive to encourage enough generic players to create price competition to rival the traditional generics market.

Thus, the insurance sector is likely to look to other solutions. One way of making treatments cheaper will be to improve medical triage, especially during the highly expensive last 6 months of life. The cold truth is that not everybody's disease is treatable. If tests can show definitively, or even on the balance of probability, that a patient won't benefit, then doctors may stop treatment and divert spending elsewhere—on palliative care, for instance. Thus, insurers will demand development of more diagnostic tests like the HER2-detecting HercepTest from Dako or the MammaPrint genetic profile from Agendia that spare breast cancer patients treatments that would be unnecessary or ineffective.

Another way to reduce disease cost will be to cut down the number of sick people. There is now a huge incentive to get behind R&D in immunization and prevention. Indeed, the legislation explicitly says that all new healthcare insurance plans will have to offer preventative care and immunizations at no cost. An integral part of any prevention program ought to be reliable and facile lifestyle diagnostics that confirm (or otherwise) patients' reporting of health-related habits such as smoking, drinking, diet and exercise—and importantly for the biotech sector, diagnostic tools that monitor and detect the early physiological signs of disease.

The need for insurers to reduce the healthcare burden also increases the incentives for developing drugs and other treatments that actually reverse disease, therapies that ameliorate the condition rather than simply treat symptoms. Disease cures—yes, we should use the phrase—are now what the US healthcare payments system will demand. And that is a real turnaround, possibly even a revolution.

IN this section



Herceptin's easy-to-use formulation p298

Pharmacy providers push for gene tests p299

Discord over new biofuel life-cycle figures p304

Safety signal dampens reception for mipomersen antisense

In February, Isis announced its cholesterol-lowering antisense therapy, mipomersen, had met its endpoints in a phase 3 trial of 124 patients with heterozygous familial hypercholesterolemia (FH) and coronary artery disease. The news ought to have sent share prices soaring, as mipomersen, if approved, could rival blockbuster statin drugs. Instead, lingering concerns related to high liver enzyme levels in some trial participants—a safety signal that had been noted in previous trials—dampened investor enthusiasm. But even if elevated liver enzymes make it less likely that mipomersen will supplant statins in major indications like coronary artery disease, efficacy data for the antisense therapy in patients with homozygous FH remain compelling.

The latest trial results provide validation that antisense drugs do have sufficient potency to rival traditional small-molecule drugs. After the clinical and regulatory disappointments of antisense therapy's other poster child, Genasense (oblimersen), Alan Gewirtz, an antisense researcher at the University of Pennsylvania School of Medicine, Philadelphia, says the results are a "shot in the arm" for the field. However, given that one touted advantage of antisense compared with other drugs is the specificity of mechanism and potentially reduced off-target effects, the elevated liver enzymes associated with mipomersen therapy are, on the one hand, disappointing. On the other, these observations are potentially informative about idiosyncratic liver toxicity specifically associated with targeting lipid biogenesis and transport.

istockphoto Donald Erickson

Mipomersen has the potential to challenge statins' market dominance, but reports of liver toxicity are causing concern.

Mipomersen is one of Isis's second-generation antisense molecules that is delivered systemically (*Nat. Biotechnol.* **25**, 497–499, 2007). The Carlsbad, California–based company is a trailblazer for oligonucleotide therapy. In 1996, it received approval for Vitravene (fomivirsen), the first antisense treatment approved by the Food and Drug

Table 1 Antisense oligonucleotides in phase 3 testing

Company/partner	Drug	Mechanism	Indication(s)
Antisense Pharma (Regensberg, Germany)	Trabedersen	A phosphorothioate antisense oligo targeting a transforming growth factor beta 2 (TGF-β2) and tumor necrosis factor ligand 13	Glioblastoma and anaplastic astrocytoma
Atlantic Healthcare (London)/Isis	Alicaforsen	2′-*O*-(2-methoxy) ethyl-modified ribose antisense oligo targeting intercellular adhesion molecule 1	Ulcerative colitis, Crohn's, inflammatory bowel disease, pouchitis, asthma
Gene Signal (Epalinges, Switzerland)	Aganirsen	A 25-mer phosphorothioate antisense oligo targeting insulin receptor substrate-1	Corneal graft rejection
Genta	Genasense (oblimersen)	A phosphorothioate antisense oligo targeting BCL-2	Melanoma, chronic lymphocytic leukemia and various other blood cancers and solid tumors
Isis Pharmaceuticals/ Genzyme	Mipomersen	2′-*O*-(2-methoxy) ethyl-modified ribose antisense oligo targeting ApoB	FH and hypercholesterolemia
NovaRx (San Diego)	Lucanix (belagenpumatucel-L)	Cells derived from four non-small cell lung cancer cell lines transfected via electroporation with a plasmid encoding a TGF-β2 antisense gene	Astrocytoma, lung tumors
OncoGeneX/Teva	Custirsin	2′-*O*-(2-methoxy) ethyl-modified ribose antisense oligo targeting clusterin	Castration-resistant prostate cancer, solid tumors

Source: Thomson Pharma Partnering; IDDB

IN brief

Bt brinjal splits Indian cabinet

AP Photo

M. Singh in pacifying mode.

India's prime minister, Manmohan Singh, has intervened in the political wrangle that erupted over a genetically modified (GM) eggplant strain due for commercial release. Approval of the locally developed *Bacillus thuringiensis* (*Bt*) variety appeared imminent, but on February 9, the minister of environment and forests, Jairam Ramesh, responded to public opposition by declaring an indefinite moratorium on the approval of *Bt* brinjal, as it is known locally, on the grounds of insufficient data to confirm that it is safe to eat. This decision has created a cabinet rift, prompting Singh to hold a consultation with senior government officials. *Bt* brinjal is India's first locally developed GM food crop and was created by Mahyco, a Jalna-based Maharashtra Hybrid Seeds Company in a joint venture with Monsanto, the St. Louis–based seed giant. Last October, the Genetic Engineering Approval Committee (GEAC), India's official regulatory body for registering GM organisms, approved release of the transgenic brinjal, opening the door for commercialization of another dozen or so GM crops in the pipeline. The environment minister's decision to over-rule the GEAC was unexpected. But it followed vociferous feedback from civil societies and advice from scientists, including Monkombu Swaminathan of 'green revolution' fame and Pushpa Bhargava, founder of the Centre for Cellular and Molecular Biology in Hyderabad, that additional safety testing of the GM brinjal was warranted. Chavali Kameswara Rao, secretary of Bangalore-based Foundation for Biotechnology Awareness and Education, believes the environment minister caved in to intense lobbying pressure from activists. He fears the resultant delay of commercialization will promote clandestine cultivation of *Bt* brinjal—similar to what happened previously in India with *Bt* cotton (*Nat. Biotechnol.* **22**, 1333–1334, 2004). On February 24, Singh and senior cabinet members agreed to set up a national biotech regulatory authority to oversee registration of transgenic products and requested that the GEAC resolve outstanding safety concerns relating to *Bt* brinjal. No time frame for these deliberations was set, however. The Department of Biotechnology (DBT), the main funding agency for transgenic research, has stayed out of the controversy. But as Prasantha Kumar Ghosh, DBT's former head of biosafety puts it, "There is no scientific basis for a moratorium." Bhim Sain Bajaj, president of the Hyderabad chapter of All-India Biotech Association, agrees. "It is a big mistake…the time will come when India will have to import rice and vegetables and we will regret [this decision]." *Killugudi Jayaraman*

Administration (FDA) for inflammation of the retina in cytomegalovirus-infected patients. Although a breakthrough experimental therapy, the treatment has not been able to build a substantial commercial market.

The company's current lead antisense product will compete for market share with the statins, a highly successful group of drugs (including Pfizer's blockbuster Lipitor; atorvastatin) that reduce levels of cholesterol, low-density lipoprotein (LDL), apolipoprotein B100 (ApoB), and triglycerides through inhibition of 3-hydroxy-3-methyl glutaryl co-enzyme A (HMG-CoA) reductase. Instead of targeting HMG-CoA reductase, mipomersen acts directly on ApoB, the protein responsible for carrying plaque-thickening LDL cholesterol into the arteries. As a second-generation 2′-*O*-(2-methoxy) ethyl-modified ribose oligonucleotide, the DNA oligonucleotide exhibits high affinity for ApoB messenger RNA. Binding of the ApoB mRNA by mipomersen then triggers cellular ribonuclease H to hydrolyze RNA phosphodiester bonds, thereby inhibiting translation and suppressing levels of ApoB protein.

On the basis of previous clinical data published last month (*The Lancet* **375**, 998–1006, 2010), mipomersen seems to do a better job than statins at fighting the dramatic cholesterol levels in individuals with FH, a rare genetic disorder. In the severe form of the disease, which affects 10,000 people worldwide, homozygous patients often have LDL levels up to six times normal, making patients susceptible to heart attacks as early as childhood. Phase 3 studies in homozygous patients led to an impressive 25% drop in LDL levels. This success prompted Isis and its partner Genzyme of Cambridge, Massachusetts, to initiate a phase 3 trial in heterozygous patients, with a view to seeking approval in the 1.5 million people who have less severe forms of FH and then extending the therapy to anyone with high cholesterol that is insufficiently controlled by statins (*Nat. Biotechnol.* **26,** 148, 2008).

The trial results announced in February for the heterozygous FH population included 124 individuals with pre-existing coronary artery disease, who were already taking maximum-tolerated doses of statins. After 26 weeks of treatment, mipomersen achieved a 28% reduction in LDL, compared with a 5% increase among controls. Isis stated that 45% dipped below 100 mg/dl—the recognized treatment goal. The study also met three secondary endpoints, with reductions in ApoB, total cholesterol and non-high-density-lipoprotein cholesterol.

The news was not all positive, however. Elevated liver enzymes were observed in 12 of 83 patients; in 5 patients, levels reached three times the upper limit of normal. It was these results, when made public, that prompted the company's shares to plummet by 17%. "Alone, [the liver enzyme increase] is not indicative of liver toxicity, but I think it serves as such a red flag, and investors have been burned so many times, they're not going to mess with this," says Edward Tenthoff, who is a senior research analyst and managing director at Piper Jaffray in Minneapolis.

In the earlier homozygous FH trial, 28 patients on mipomersen experienced a 24.7% decrease in LDL, compared with a 3.3% reduction among the placebo group. Four patients (12%) had increases in concentrations of the liver enzyme alanine aminotransferase (ALT) of at least three times the upper limits of the normal range. Other liver tests, including levels of bilirubin, albumin and prothrombin, showed no signs of liver damage, according to the company.

Importantly for antisense technology as a whole, the spike in enzyme levels is not entirely unexpected; in fact, it is likely to be target specific, given that lipid-lowering drugs are known to be one of the rare examples in which a pharmacodynamic property of the drug class as a whole accounts for liver toxicity. Statins, for example, are associated with a dose-related increase in the incidence of the liver ALT three times greater than the upper limits of the normal (threshold set by regulators); indeed, acute liver failure occurs in about one in a million statin-treated patients. The 124 patients in the mipomersen trial were already on high doses of statins.

For its part, Isis believes that elevated liver enzymes are not an inherent problem of antisense technology. For one thing, individuals who had the steepest drop in LDL and ApoB levels on mipomersen tended to have higher enzyme levels. What's more, ~5,000 people have been treated with oligonucleotides against other targets, with no evidence of ALT increases, according to Geary. "[The elevated enzymes] are more likely an on-target side effect that's related to the mechanism. I don't know that this is necessarily a black mark on antisense," says Brian Abrahams, senior biotech analyst at Oppenheimer & Co in New York.

As yet, the mechanism underlying the increase in liver enzymes in a small number of patients is not well understood. Possible explanations include alterations in the lipid concentrations in hepatocyte membranes, which could cause mild hepatic cell dysfunction, "but as far as I can tell it's uncer-

tain," says Cy Stein, professor of medicine and molecular pharmacology at the Albert Einstein College of Medicine in New York. By acting on ApoB, a carrier for lipids, mipomersen may also affect fat accumulation in the liver. Suppression of ApoB may lead to the accumulation of fats in cellular lipids, which might trigger the raised liver enzyme levels observed. This idea is supported by some patients with a genetic condition called hypobetalipoproteinemia, who cannot make ApoB and have low levels of LDL. Some of these patients accumulate liver fat. Thus, mipomersen could, in effect, be mimicking that condition, says Robert Hegele, a professor of medicine and biochemistry and director of the Blackburn Cardiovascular Genetics Laboratory at Robarts Research Institute in London, Ontario. Still, "It's all speculation [at this point]," he hastens to add.

Antisense therapies as a class, on the other hand, do have off-target effects of their own. One issue is that the highly charged phosphorothioate backbone binds tightly to charged residues in proteins—a property that helps them avoid elimination via the kidney through association with albumin proteins. But this property might also lead to the binding of antisense to proteins on the surface of hepatocytes, perhaps mimicking heparin and leading to abnormalities, according to Stein.

Despite the recent setback, Genzyme management, which is partnered with Isis on mipomersen, remains sanguine, noting that the effects on liver enzymes were reversible. "Physicians can manage (side effects) by backing off the medication because they're seeing such a significant drop in LDL," says Paula Soteropoulos, vice president and general manager of Genzyme's cardiovascular business. But some analysts think it's likely that mipomersen will be approved only for homozygous FH patients, who cannot metabolize LDL due to a lack of functional LDL receptors responsible for clearing LDL from plasma. Even if it were approved for a broader population, mipomersen might not have sufficient advantages to convince physicians to switch from small-molecule statin therapies that are administered orally rather than subcutaneously injected.

In any case, Genzyme plans to file in the first half of 2011 in the US and Europe, targeting patients with homozygous FH and possibly severe hypercholesterolemia. Both indications together represent ~25,000 patients in the US and Europe.

The impressive efficacy of mipomersen is testament to Isis's investment and optimization of second-generation antisense technology. These second-generation chemistries improve stability and binding, Soteropoulos says, and Isis has spun off or licensed the technology to other companies, such as OncoGeneX, located in Bothell, Washington (**Table 1**), and Altair Therapeutics, of San Diego.

Not everyone is convinced that a rejuvenation in antisense approaches is on the horizon, however. "Isis is the only company left. Other companies have converted to CpG or siRNA [small interfering RNA] approaches," says John Rossi, a professor of molecular and cellular biology at the City of Hope's Beckman Research Institute, Duarte, California. The main advantage of RNA interference (RNAi) over antisense has been its greater potency. "RNAi is long lasting. Once it's engaged the RNA silencing complex, [siRNA] can last for weeks," says Rossi. University of Pennsylvania researcher Gewirtz, agrees: "It just seems easier to find an RNA molecule that gets you into the game than it is to find an oligo. That's why RNAi became so widely accepted—it just works for everybody."

Even so, siRNA and CpG suffers from the same issues of off-target effects and delivery as antisense, and the latter has other advantages. Its easier to manufacture than siRNA, and because antisense has been around longer, there is more clinical experience behind it. There have been 20 or so clinical studies involving antisense, according to Tenthoff, whereas RNAi trials are still in the single digits. "I don't think antisense is yet giving way to RNAi. Antisense is still a more clinically experienced technology," says Tenthoff.

"I'm very enthusiastic about both approaches," says Raymond P. Warrell Jr., CEO of Genta. "The advantage of antisense is that there are now 15-plus years of clinical experience with it. The folks working primarily on RNAi are in the process of relearning a lot of [those] lessons." He expects the two technologies to ultimately be complementary. "Whether you use RNA or DNA depends to some extent on the target and to some extent the technique you have experience with. The challenge remains first and foremost to identify a critical target so you have a high level of confidence that knocking it out or down will have a transformative effect on the disease," says Warrell.

Jim Kling *Bellingham, Washington*

IN brief

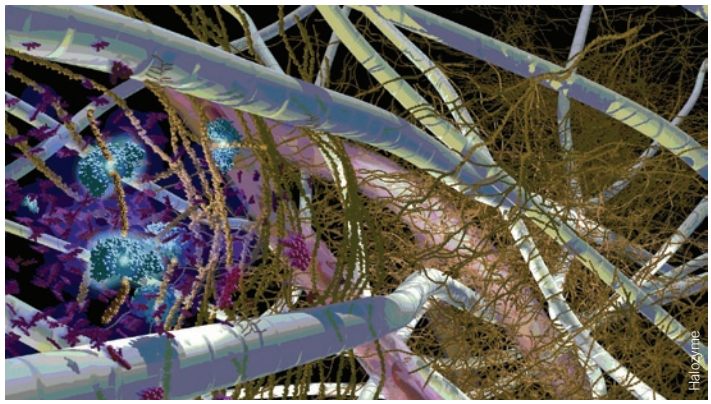
Orphans on the rise

The number of drug approvals for orphan indications has doubled in recent years, according to a report from the Tufts Center for the Study of Drug Development. The independent, nonprofit research group at Tufts University in Boston, found that between 2000 and 2002 the US Food and Drug Administration (FDA) approved 208 orphan drugs, and the number climbed to 425 between 2006 and 2008. The increase could reflect the fact that orphan diseases are simple to target as they are often underpinned by a single genetic cause. But financial incentives for pursuing orphan drugs, such as a waiver of the FDA's $1.4 million filing fee, long marketing exclusivity and high prices charged are likely factors, too. According to FDA data, biotech firms generate 50% of orphan drug applications and academia another 25%. Pharma makes up less than 25% of the total probably because "orphan drugs do not regularly fit their business model," says Tim Coté, director of the FDA's Office of Orphan Product Development. That has been changing, however. Regeneron, of Tarrytown, New York, has seen this firsthand with Arcalyst (rilonacept), an interleukin-1 (IL-1) 'trap' to treat cryopyrin-associated periodic syndromes (CAPS), a rare disease that affects only a few thousand people globally. In 2003, Novartis of Basel terminated a collaboration with Regeneron over the IL-1 trap, because the Swiss pharma was not interested in the small market for CAPS. Later, Novartis developed an IL-1 antibody for CAPS, Ilaris (canakinumab), which gained approval in July 2009. *Nadine Kolas*

Pharma's Asian syndicate

Three big pharmas—Pfizer, Merck, and Eli Lilly—are pooling their resources to set up an independent nonprofit company to spur research into innovative treatments for cancers common in Asian populations. The new Asian Cancer Research Group (ACRG) will build an open-access pharmacogenomic cancer database, which will be made publicly available to researchers in the field. Wu Jun, vice president of Xiangxue Pharmaceutical, Guangzhou, says, "It will save Western companies time and money and is good news for patients in China." The joint venture will focus initially on lung and gastric cancers and aims to gather 2,000 tissue samples over the next two years. "ACRG could get more data from Asia and spend less on research compared with what they spend in the West," says Wu. ACRG is an example of a growing trend in pre-competitive collaborations. The same three companies have done it before with Enlight Biosciences (*Nat. Biotechnol.* **26**, 960–961, 2008), an R&D startup for developing drug discovery tools. The pharma giants are searching for ways to capture the emerging Asian markets. Plans for ACRG were already underway before last year's decision by the Chinese government to invest 850 billion yuan ($125 billion) on healthcare reform, according to a spokesperson for Merck. *Bea Perks*

Roche plans for more convenient-to-use Herceptin and Rituxan



The extracellular matrix is temporarily broken down by Halozyme's human hyaluronidases, opening an opportunistic route to deliver biologics subcutaneously.

In January, the Swiss drug manufacturer Roche announced a €190 million ($175 million) investment to manufacture a device that enables the subcutaneous injection of biologics, notably its blockbuster antibody cancer drug Herceptin (trastuzumab), by combining it with an enzyme that opens up channels in the extracellular matrix. San Diego–based Halozyme Therapeutics is supplying the enzyme, a recombinant hyaluronidase it has developed. The new Herceptin formulation is now in phase 3 clinical trials, and if the program were successful, it would offer patients a convenient dosing form of the drug, which is currently given via intravenous (i.v.) infusion. As a line extension, it expands the blockbuster drug franchise and extends the patent protection for Herceptin—a strategy Roche could apply to other biologics as well.

"When we looked at the Roche/Genentech portfolio of biologics, it became quite clear that, if we wanted a more convenient administration form, we needed tools that allow us to painlessly inject larger volumes," says Johannes Schmidt, project leader, biologics, for Roche in Basel. Roche completed its acquisition of Herceptin's developer, Genentech, in San Francisco, in 2009. Before that, the Swiss pharma owned part of the company and held marketing rights outside the US to Herceptin as well as other Genentech drugs.

Typically, the volume of a subcutaneous injection is limited to around 1 ml. Larger injection volumes are painful and cause tissue distortion, edema and irritation and redness (erythema). That's not an issue with some biologics, such as the tumor necrosis factor alpha inhibitors used to treat rheumatoid arthritis, where the volumes involved are very small. But the situation is different with cancer drugs. Herceptin, for example, is given at doses of 500 mg or more. "A concentrated antibody formulation like 100 mg/ml would take at least 5 ml to inject someone," says Schmidt. "That would not be possible without real pain."

Roche turned to Halozyme to solve the problem. The companies initiated a collaboration around recombinant human hyaluronidase (rHuPH20) in 2006. Hyaluronidases degrade hyaluronic acid (hyaluran) in the extracellular matrix, creating channels about 200 nm in diameter through which large molecules can pass to get into the bloodstream. "Most biologics, including monoclonal antibodies, cytokines, recombinant enzymes, other proteins and peptides can be enabled by the technology," says Halozyme CEO Jonathan Lim. Bovine hyaluronidase has been in use for >50 years, mostly for ophthalmic surgery, where the enzyme facilitates local distribution of anesthetics. But repeat administrations of an animal-derived molecule cause immunogenicity, making it inappropriate for cancer maintenance therapy.

Recombinant HuPH20 suffers no such drawbacks, and it is also the only enzyme in the hyaluronidase family that works at neutral pH and has only temporary activity. "You only want to open the subcutaneous tissue to generate a cavity for a small period of time," says Schmidt. "You don't want to leave gaps there." With rHuPH20, the gap is closed in 24 hours.

In addition to its collaboration with Roche, Halozyme has licensed rHuPH20 to Baxter Healthcare, in Northbrook, Illinois, for use in its Gammagard plasma-derived immune globulin. Halozyme is also developing a very fast-acting insulin by combining it with rHuPH20 to create a more physiologic pharmacokinetic profile that potentially allows for better glycemic control and potentially less weight gain. But Roche is not attempting to change the pharmacokinetics of its drugs using rHuPH20. Indeed, having a more convenient formulation otherwise comparable to the i.v. version helps speed the regulatory pathway.

Roche has brought three rHuPH20-formulated products into clinical trials with already-approved drugs, including Herceptin, the leukemia and lymphoma antibody MabThera (rituximab, Rituxan in the US), and an undisclosed third candidate; the latter two are each in phase 1 safety studies. Initial results of the phase 3 Herceptin trial, which began dosing patients last October, are expected toward the end of 2011, at which time Roche could seek marketing approval. "We can rely on the huge safety database of Herceptin and bridge to the Herceptin i.v. data," Schmidt says.

Herceptin was the most obvious candidate to establish proof of concept—especially for the part of the treatment where it is administered as monotherapy, Schmidt adds. "That is mainly the case in early breast cancer, once the patient has finished chemotherapy." Roche is also studying the efficacy of using Herceptin for two years instead of one, as maintenance therapy. "If two-year comes out to be superior, we have to have a more convenient regimen to support the patients," he says. Subcutaneous delivery would also have compliance benefits. As an infusion, "we have seen that patients don't even complete their one-year [program] sometimes," he says. "If they go to a clinic and see sick patients all around them, they may actually fall back, because in early breast cancer maintenance treatment they don't feel sick."

The drug is delivered subcutaneously as a single premixed solution containing rHuPH20, which is built into a glass cartridge that is packed into a motor-driven device. "The patient essentially only has to pull the cover of the adhesive off," says Schmidt. "Everything else is done automatically." The needle is inserted, the device injects the solution and the needle automatically retracts. In some countries (e.g., Germany), the device can be self-administered, whereas in others it will be used in the physician's office, depending on the country's needle requirements. The device would potentially work with any stable liquid formulation of a biologic. "Trastuzumab and rituximab are fairly stable antibodies," says Schmidt. "There may be others where that's an issue and we probably couldn't use our injection device."

Roche's production line in Kaiseraugst, Switzerland, will supply material for clinical studies and market launch of subcutaneous Herceptin, whereas its full-scale automated production line in Mannheim, Germany will deliver commercial supply to markets. Roche told investors in February that given the dynamics of the US oncology market, where physicians derive income from i.v. infusions, it is focusing on opportunities for subcutaneous delivery outside the US.

Mark Ratner *Cambridge, Massachusetts*

US pharmacies broaden access to pharmacogenetic tests

Medco Health Solutions, the pharmacy services manager, bolstered its commitment to personalized medicine in February, with the acquisition of DNA Direct, a San Francisco–based genomic medicine company. The terms of the deal were not disclosed, but Medco of Franklin Lakes, New Jersey, has now incorporated more than 2,000 genetic and molecular tests into its books. The aim is to become a one-stop health service shop aimed at helping physicians and payors better match individuals to therapeutics and improve clinical outcomes and save money. But whether personalized medicine can actually cut costs, and whether pharma companies will remain on the sidelines now that pharmacy services firms are pushing to control when and how drugs are used, is yet to be seen.

Medco is not the first pharmacy-service firm to close a deal in the pharmacogenetics area. In November, pharmacy healthcare provider CVS Caremark of Woonsocket, Rhode Island, signed a strategic partnership with Generation Health located in Upper Saddle River, New Jersey. Both pharmacy benefits managers had already set up some personalized medicine services, particularly in oncology. Medco, for instance, established a personalized medicine research center headed by Felix Frueh, formerly a specialist on personalized medicine at the US Food and Drug Administration (FDA).

Several trends helped draw pharmacy-benefit companies into the personalized medicine arena. One is the huge explosion in the number of genetic tests. Medco's Jane Barlow estimates that "up to a 1,000 [current genetic tests] need more benefits management around them." She says there is about 20% waste in the testing business because of inappropriate interpretation or just because doctors aren't sure how to use new tests (*Nat. Biotechnol.* **28**, 117–119, 2010). "This is an area that continues to grow rapidly and the information changes overnight," says Barlow, who is vice president of medical strategy and clinical quality at the giant pharmacy provider.

Another factor is growing interest from payors. "To our clients, the insurance companies and self-insured employers, being able to provide tests that target drugs to individuals is of great interest," says Troyen Brennan, chief medical officer at CVS Caremark. At the moment, the "lion's share" of medicines with the potential to be personalized is in oncology, and most of these drugs are administered in physicians' offices. But Brennan sees that moving rapidly to medications dispensed at pharmacies and by mail order.

Graham Gardner is both a venture capitalist at Highland Capital Partners, in Lexington, Massachusetts, and the chief medical officer of genetic-testing benefits management firm Generation Health, which Highland helped start up in 2008. Gardner thinks that consumers are now more interested and aware of genetic testing than in the past thanks to personal genomics companies, such as the California-based Navigenics located in Foster City, and 23andMe, headquartered in Mountain View.

The fact that this is happening during the height of uncertainty about health reform makes

Medco Research Institute

Medco is encouraging doctors to screen patients using genetic tests before prescribing blood thinner warfarin and breast cancer drug tamoxifen.

IN brief

RNAi patent jolt

The US Patent and Trademark Office has issued a patent for detection of RNA-mediated gene silencing to Sir David Baulcombe, University of Cambridge, and Andrew Hamilton, University of Glasgow, over a decade after their gene silencing findings in plants were first reported (*Science* **286**, 950–952, 1999). "The new patent has implications beyond plants," says Jan Chojecki, CEO of Plant Bioscience Limited (PBL), of Norwich, the tech transfer company that owns the patents. "Anyone in the US profiling short RNAs and their impact on gene expression in mammalian systems is likely to be interested. We think it will create quite a stir." The new patent recognizes Baulcombe and Hamilton's discovery that when genes are silenced complementary RNA strands of 20–30 bp accumulate—a finding that also proved critical to establishing short RNAs as a tool to manipulate gene expression. The initial patent for this technology, issued in 2004, was limited to plants, but the new patent broadens out to mammals. PBL expects to grant licenses to industry but will not enforce rights in academia, provided researchers use licensed detection kits. James McNamara, who directs the Office of Technology Management, University of Massachusetts Medical School, points out that Craig Mello and Andrew Fire, now at Stanford University, developed comparable RNA detection methods. "But if a company practices methods that might infringe on Baulcombe and Hamilton, they might take a license on it for reasonable terms," he says. *Charlie Schmidt*

Court voids HGS gene patent

In the first British case to deal with the validity of a gene sequence patent, a UK Court of Appeal struck down a patent held by Rockville, Maryland–based Human Genome Sciences (HGS) for lack of industrial application. The dispute in *Eli Lilly & Company v Human Genome Sciences, Inc.* centers around the validity of a patent for the neutrokine-α gene sequence. The outcome agreed on in February is expected to have a major bearing on future decisions on the scope of biotech patents, especially gene sequence patents. "The Court of Appeal's judgment will raise doubts over the validity of many biotech patents currently in force, as mere speculation [on] biological function—without experimental data—may not be sufficient, says intellectual property expert Robert Fitt at London-based law firm Bristows. "With pharma's increasing reliance on biotech drugs as a source of growth, the value of many patent portfolios may well be hit hard by this judgment," he says. HGS's patent was initially struck down by the UK Patents Court in July 2008 but later held valid by the Technical Board of Appeal of the European Patent Office. HGS is partnered with GlaxoSmithKline of London to develop an antibody to neutrokine-α called Benlysta (belimumab), for lupus. Lilly is developing its own antibody to neutrokine-α, having already spent some $50 million on its development, with plans to spend another $250 million in clinical trials. *Michael Francisco*

it all the more interesting. Some experts point to the impending comparative effectiveness package as one reason for the surge in personalized medicine approaches. Others are skeptical. "It takes as long as ten years for countries to set up comparative clinical effectiveness platforms," says Paul Keckley, executive director of the Deloitte Center for Health Solutions. "At the end of the decade, that is a strategic opportunity for personalized medicine, but it's a long way away."

Cost reduction is probably the main driver, Generation Health's Gardner believes. "[Genetic testing] addresses one of the areas where costs are growing the fastest and are still largely unmanaged," he says. Keckley agrees that the need to cut healthcare costs is most likely behind the surge in interest in personalized medicine. "Next-generation diagnostics, like Genomic Health's OncoDx, stand to profit from that trend, not drug companies making personalized medicines," he says.

Venture capitalist Dion Madsen and his colleague Stacy Feld at Physic Ventures, San Francisco, also see this as being a moment for diagnostics companies, not pharma. "We're interested in platforms and tools that will enable and inform decision making in the doctor's office or at the hospital," says Feld. The point-of-care aspect is critical in their view. "And we take a broader view of that," Madsen explains. "The point-of-care could be a retail clinic or the consumer themselves." In keeping with that strategy, the group recently invested in the personalized diagnostic firm On-Q-Ity, of Waltham, Massachusetts, which is developing a microfluidic device capable of detecting circulating tumor cells (based on work by Mehmet Toner's group at Harvard-MIT) and raised $26 million in a series A funding round.

The new focus on genetic tests will put drug companies in a predicament, says Peter Keeling, CEO of personalized medicine consulting firm, Diaceutics (Belfast, Northern Ireland). "This is a game changer for pharma, because now someone else will be determining how their drugs are used," Keeling says. Because they will control the testing and apparently much of the growing research in this field, the PBMs [pharmacy benefits managers] will start establishing which patients get which drugs, not the drug companies, as has traditionally been the case. Keeling also sees growing interest in personalized medicine on the part of payors. "A couple years ago they were all sitting on the fence," he says. "That's changed."

Nobody's expecting a stampede of pharmacy benefits managers or payors to follow Medco and CVS Caremark. "I think these companies are making these moves as part of their long-term strategies," says Philip Ma, director in McKinsey & Company's Silicon Valley office and the leader of the company's West Coast Healthcare Practice. "There is no pressing reason for them to enter this market now." Ma points out that there is a lot more "low hanging fruit" for them to squeeze money out of before they'll get noticeable savings from personalized medicine. "I expect they are a lot more focused on the basics, such as negotiating for lower drug prices, getting doctors and patients to comply with formularies, and working with physicians to control use," he says.

But optimism surrounding the feasibility of personalized medicine has been growing. In December, PricewaterhouseCoopers released a report estimating that the core market for personalized medicine—diagnostics and therapeutics—is already worth $24 billion and expected to grow by 10% annually, reaching $42 billion by 2015.

In the short term, however, McKinsey's Ma sees many hurdles to the personalized approach. Doctors may not have many reasons to change their practices yet, and patients can always switch doctors if they feel they are being denied a drug that might help them. Meanwhile, there is still tremendous uncertainty around regulatory issues. Personalized medicine is often spoken of by FDA staffers, but progress on critical issues, such as a drug and companion diagnostic approval pathway, has been very slow. As *Nature Biotechnology* went to press, the FDA's Pharmaceutical Science and Clinical Pharmacology Advisory Committee was meeting to discuss application of pharmacogenomics in the early stages of drug development. FDA commissioner Margaret Hamburg also recently spoke at a luncheon sponsored by the nonprofit, Washington, DC–based Personalized Medicine Coalition. According to a transcript, Hamburg said that the FDA would "issue draft process guidance on biomarker qualification" in the next few months. She also said, "We intend to clarify our expectations for the kinds of clinical trials and levels of confidence needed to satisfy us that a test is accurate and that it can be used to help shape clinical judgments."

If the FDA does take those steps, it might help convince a few more investors that it's time to jump on the personalized medicine bandwagon or risk missing the action. Keckley is optimistic, pointing to the fact that the FDA approved 26 drugs in 2009 and six of those are personalized medicines. "I think Margaret [Hamburg] inherited a fairly dysfunctional and inefficient operation, and she's already improved it," he says. The other missing ingredient, for investors at least, is knowing what direction healthcare reform will go. "There is tremendous hesitancy to move in any direction as long as there is such a big question mark about reform," Keckley says.

Malorye Allison *Acton, Massachusetts*

IN brief

RNAi patent jolt

The US Patent and Trademark Office has issued a patent for detection of RNA-mediated gene silencing to Sir David Baulcombe, University of Cambridge, and Andrew Hamilton, University of Glasgow, over a decade after their gene silencing findings in plants were first reported (*Science* **286**, 950–952, 1999). "The new patent has implications beyond plants," says Jan Chojecki, CEO of Plant Bioscience Limited (PBL), of Norwich, the tech transfer company that owns the patents. "Anyone in the US profiling short RNAs and their impact on gene expression in mammalian systems is likely to be interested. We think it will create quite a stir." The new patent recognizes Baulcombe and Hamilton's discovery that when genes are silenced complementary RNA strands of 20–30 bp accumulate—a finding that also proved critical to establishing short RNAs as a tool to manipulate gene expression. The initial patent for this technology, issued in 2004, was limited to plants, but the new patent broadens out to mammals. PBL expects to grant licenses to industry but will not enforce rights in academia, provided researchers use licensed detection kits. James McNamara, who directs the Office of Technology Management, University of Massachusetts Medical School, points out that Craig Mello and Andrew Fire, now at Stanford University, developed comparable RNA detection methods. "But if a company practices methods that might infringe on Baulcombe and Hamilton, they might take a license on it for reasonable terms," he says. *Charlie Schmidt*

Court voids HGS gene patent

In the first British case to deal with the validity of a gene sequence patent, a UK Court of Appeal struck down a patent held by Rockville, Maryland–based Human Genome Sciences (HGS) for lack of industrial application. The dispute in *Eli Lilly & Company v Human Genome Sciences, Inc.* centers around the validity of a patent for the neutrokine-α gene sequence. The outcome agreed on in February is expected to have a major bearing on future decisions on the scope of biotech patents, especially gene sequence patents. "The Court of Appeal's judgment will raise doubts over the validity of many biotech patents currently in force, as mere speculation [on] biological function—without experimental data—may not be sufficient, says intellectual property expert Robert Fitt at London-based law firm Bristows. "With pharma's increasing reliance on biotech drugs as a source of growth, the value of many patent portfolios may well be hit hard by this judgment," he says. HGS's patent was initially struck down by the UK Patents Court in July 2008 but later held valid by the Technical Board of Appeal of the European Patent Office. HGS is partnered with GlaxoSmithKline of London to develop an antibody to neutrokine-α called Benlysta (belimumab), for lupus. Lilly is developing its own antibody to neutrokine-α, having already spent some $50 million on its development, with plans to spend another $250 million in clinical trials. *Michael Francisco*

it all the more interesting. Some experts point to the impending comparative effectiveness package as one reason for the surge in personalized medicine approaches. Others are skeptical. "It takes as long as ten years for countries to set up comparative clinical effectiveness platforms," says Paul Keckley, executive director of the Deloitte Center for Health Solutions. "At the end of the decade, that is a strategic opportunity for personalized medicine, but it's a long way away."

Cost reduction is probably the main driver, Generation Health's Gardner believes. "[Genetic testing] addresses one of the areas where costs are growing the fastest and are still largely unmanaged," he says. Keckley agrees that the need to cut healthcare costs is most likely behind the surge in interest in personalized medicine. "Next-generation diagnostics, like Genomic Health's OncoDx, stand to profit from that trend, not drug companies making personalized medicines," he says.

Venture capitalist Dion Madsen and his colleague Stacy Feld at Physic Ventures, San Francisco, also see this as being a moment for diagnostics companies, not pharma. "We're interested in platforms and tools that will enable and inform decision making in the doctor's office or at the hospital," says Feld. The point-of-care aspect is critical in their view. "And we take a broader view of that," Madsen explains. "The point-of-care could be a retail clinic or the consumer themselves." In keeping with that strategy, the group recently invested in the personalized diagnostic firm On-Q-Ity, of Waltham, Massachusetts, which is developing a microfluidic device capable of detecting circulating tumor cells (based on work by Mehmet Toner's group at Harvard-MIT) and raised $26 million in a series A funding round.

The new focus on genetic tests will put drug companies in a predicament, says Peter Keeling, CEO of personalized medicine consulting firm, Diaceutics (Belfast, Northern Ireland). "This is a game changer for pharma, because now someone else will be determining how their drugs are used," Keeling says. Because they will control the testing and apparently much of the growing research in this field, the PBMs [pharmacy benefits managers] will start establishing which patients get which drugs, not the drug companies, as has traditionally been the case. Keeling also sees growing interest in personalized medicine on the part of payors. "A couple years ago they were all sitting on the fence," he says. "That's changed."

Nobody's expecting a stampede of pharmacy benefits managers or payors to follow Medco and CVS Caremark. "I think these companies are making these moves as part of their long-term strategies," says Philip Ma, director in McKinsey & Company's Silicon Valley office and the leader of the company's West Coast Healthcare Practice. "There is no pressing reason for them to enter this market now." Ma points out that there is a lot more "low hanging fruit" for them to squeeze money out of before they'll get noticeable savings from personalized medicine. "I expect they are a lot more focused on the basics, such as negotiating for lower drug prices, getting doctors and patients to comply with formularies, and working with physicians to control use," he says.

But optimism surrounding the feasibility of personalized medicine has been growing. In December, PricewaterhouseCoopers released a report estimating that the core market for personalized medicine—diagnostics and therapeutics—is already worth $24 billion and expected to grow by 10% annually, reaching $42 billion by 2015.

In the short term, however, McKinsey's Ma sees many hurdles to the personalized approach. Doctors may not have many reasons to change their practices yet, and patients can always switch doctors if they feel they are being denied a drug that might help them. Meanwhile, there is still tremendous uncertainty around regulatory issues. Personalized medicine is often spoken of by FDA staffers, but progress on critical issues, such as a drug and companion diagnostic approval pathway, has been very slow. As *Nature Biotechnology* went to press, the FDA's Pharmaceutical Science and Clinical Pharmacology Advisory Committee was meeting to discuss application of pharmacogenomics in the early stages of drug development. FDA commissioner Margaret Hamburg also recently spoke at a luncheon sponsored by the nonprofit, Washington, DC–based Personalized Medicine Coalition. According to a transcript, Hamburg said that the FDA would "issue draft process guidance on biomarker qualification" in the next few months. She also said, "We intend to clarify our expectations for the kinds of clinical trials and levels of confidence needed to satisfy us that a test is accurate and that it can be used to help shape clinical judgments."

If the FDA does take those steps, it might help convince a few more investors that it's time to jump on the personalized medicine bandwagon or risk missing the action. Keckley is optimistic, pointing to the fact that the FDA approved 26 drugs in 2009 and six of those are personalized medicines. "I think Margaret [Hamburg] inherited a fairly dysfunctional and inefficient operation, and she's already improved it," he says. The other missing ingredient, for investors at least, is knowing what direction healthcare reform will go. "There is tremendous hesitancy to move in any direction as long as there is such a big question mark about reform," Keckley says.

Malorye Allison *Acton, Massachusetts*

Public companies get creative in raising finance

Ablynx, whose Belgium headquarters are pictured above, is the latest publicly quoted biotech in Europe to strike a strategic financing deal.

The beginning of March saw that rarest of events, a substantial secondary public offering of stock from a European biotech company. Ablynx, of Ghent, Belgium, raised €50 million to fund a range of new and continuing clinical programs of its single-domain antibody fragment (or 'nanobody') products. Although CEO Edwin Moses describes the offering as taking Ablynx "to the next level in corporate development," others see the move as a sign that the public markets are offering biotech companies more flexibility in raising finance than in the past. At the same time, several firms are embracing a new type of hybrid financing approach—the standard equity distribution agreement (SEDA)—which offers companies advantages over previous refinancing options.

Whereas secondary offerings (secondaries) have fueled much of the growth of the US biotech sector, European companies have consistently struggled to execute them. In the 1990s, a few UK companies, such as Oxford-based British Biotech and Oxford GlycoSystems, managed to raise $100-million-plus secondaries. Unfortunately, the demise of these companies discouraged institutional investors, and subsequent initial public offerings (IPOs), both in the UK and mainland Europe, tended to be tightly controlled events in which existing venture capital investors retained a large proportion of the shares, a situation that left potential institutional investors reticent to buy into European biotech offerings—either IPOs or secondaries.

What is remarkable about the Ablynx offering is not its size *per se* but its size relative to the volume of share dealing. In the US last year, companies like Dendreon, Vertex and Rockville, Maryland–based Human Genome Sciences, each completed multi-hundred-million-dollar secondary offerings. Seattle-headquartered Dendreon's secondary offering last May raised a whopping $200 million, but this represents just twice the value of Dendreon shares that change hands every day on NASDAQ. For Vertex, located in Cambridge, Massachusetts, its $477 million secondary offering in December works out at just 5 or 6 times its daily trading value. Ablynx's €50 million, on the other hand, represents around 250 times its trading value, because only around 25,000 Ablynx shares change hands each day. A similar multiplier applies to another Belgian company, Movetis, located in Antwerp, which completed its €85 million IPO last December. Possibly top of the heap in this regard, however, is London-based Proximagen, a neuroscience specialist, which raised £50 million in a secondary offering in June 2009 and yet trades less than 1,000 shares per day on the UK's Alternative Investment Market at a value of £1.00–1.25. Its ratio of money raised to daily trade value is >50,000!

The lack of liquidity for public biotechs, particularly those listed on European exchanges, has meant that after flotation they have been dependent on their venture investors or corporate partners (such as pharma companies) to top up financing over long periods. In March, for instance, Archimedes Pharma, of Reading, UK—a company that has a portfolio of niche medicines already on the market and is widely expected to float in 2010—had to return to venture investors, Novo Growth Equity and Warburg Pincus, for £65 million ($100 million).

In this context, the challenge for Ablynx has been to find enough funding to build a portfolio of clinical products from its 'nanobody' platform technology, a form of single-variable-domain antibodies. Financing such a transition requires substantial capital. Since its foundation in 2001, the company has raised over €70 million from venture capitalists, €85 million in an IPO on Brussels EuroNext in November 2007 and ~€70 million from its corporate partnerships. However, many of those partnership deals were signed at a very early stage, meaning that Ablynx can reap only limited benefits from the products.

The dilemma for the company, of course, is that any participation in clinical development is very capital intensive. This is why Ablynx, which reported cash reserves of around €90 million at the end of its 2009 financial year, decided to go back to the market. "We would now like to take programs past proof of concept," says Ablynx's Moses, and that means that the €50 million plus raised in the secondary offering has to work very hard. The company predicts that its first product will hit the market in 2013. "We think that we could eke out the money until then," says Moses, "but this is a worst-case scenario. In the best case, we will demonstrate that we can take product to the clinic on our own and develop value from the platform; and that will enable us either to strike deals that are more valuable or go back to the [money] markets."

Ablynx's market move illustrates how publicly quoted biotech companies in Europe need to take a more strategic approach to financing than they have so far. According to Dennis Purcell, senior managing partner at Aisling Capital of New York, the significant secondary offerings in the US last year—Vertex, Dendreon and Human Genome Sciences' $477 million in December—have all been associated with pipeline commercialization. For all three, the triggers have been the imminent appearance of wholly owned or partnered products on the market. "Institutional investors [upon whom the magnitude of secondary offerings depends] only want to invest when the company is at an inflexion point," says Purcell, "and the inflexion point these days tends to be product approval rather than the emergence of clinical data."

But most European public companies do not have the luxury of imminent product approval. For firms that are public companies already, further capital raising has depended

Table 1 Biotech and specialty pharma companies involved in SEDAs

Company	Location	Provision	Start
Advanced Life Science	Woodridge, Illinois	$15 million	September 2008
Newron Pharmaceuticals	Milan, Italy	30 million CHF	December 2008
MediGene	Martinsried, Germany	€25 million	December 2008
RXI Pharmaceuticals	Worcester, Massachusetts	$25 million	January 2009
Pharming	Leiden, The Netherlands	€20 million	April 2009
Achillion Pharmaceuticals	New Haven, Connecticut	$15 million	July 2009
Hybrigenics	Paris	€15 million	September 2009
Sygnis Pharma	Heidelberg, Germany	€10 million	September 2009
Biotie	Helsinki, Finland	€20 million	October 2009
Labopharm	Laval, Quebec, Canada	$20 million	November 2009
Allon Therapeutics	Vancouver, British Columbia, Canada	$10 million	March 2010

Source: Yorkville Associates

largely on the PIPE (private investment in public equity) or 'PIPE-like' instruments. In a PIPE, a range of private institutional investors are offered newly issued stock at a discount to the current market rates. The problem with the PIPE, according to Thomas Klaue, CFO at MediGene, Martinsried, Germany, is that activity in the stock market by hedge funds and speculators tends to put pressure on a company's stock price. "All the arbitrage players bet against your stock behavior," says Klaue, "and this makes it an expensive way to raise money."

This is why Klaue is an enthusiast for a hybrid financing instrument called the SEDA offered by Yorkville Advisors of Jersey City, New Jersey. In executing a SEDA, Yorkville acts both as a conventional underwriting bank and an investor. Like an underwriter, it makes a commitment to buy a certain amount of stock at a price based on the prevailing stock price. The big difference from merchant banking behavior in a conventional offering, though, is that Yorkville doesn't simply try to sell the stock straightaway and return the yield to the company. Instead, it holds the stock and waits for a better price. Instead of the traditional banker's fee, Yorkville has to earn its way in the world with the discount of around 5% it receives on the stock price. Its upside is dependent on the strengthening stock price.

For Thomas Klaue, this deal structure better aligns the interests of investors and the corporation. "In a traditional offering, if the underwriter dumps the stock on the market and the stock price goes down, only the company loses yield. The underwriter still earns its fee," he explains. MediGene put its SEDA provision in place in December 2008 but only drew down on it in the fourth quarter of 2009. "The SEDA is like a small PIPE that you can call at your own discretion. As CFO, I have access to money without actually having the money right now. And the fund can ease the stock into the market, keeping the price strong—which is in their interest."

Yorkville's managing director of healthcare, Michael Nowak, who has been responsible for putting in place some 11 SEDAs since the end of 2008 (**Table 1**), points out that Yorkville has an obligation to buy stock at any time specified by the company, but conversely the company has no obligation to sell even a single share. Or it can wait until a "good news" story drives up the stock price and ask Yorkville to buy into what may prove to be a bubble. "The SEDA is a potential source of funding," he explains, "with a right but not an obligation on the company to draw down all or any of the money." At least in theory, it gives a company leverage to go into business development or other financing discussions in a financially more secure position and negotiate a better deal. "I am fine with leveraging by the company," says Michael Nowak. "We want to be an investor in good fundamentals, and if the existence of the SEDA is used to further the development of a company, then that is a sign of good management."

Nowak sees the use of instruments like SEDA as a sign of market maturity. "Companies get a chance to do something different," he says. "They have a flexible option that they can use to navigate through the shoals, and they don't have to raise all their capital at today's price, which may be low and therefore very dilutive."

John Hodgson *Cambridge, UK*

IN their words

"You load a tractor trailer with drugs and it's worth more than refrigerators." After the latest heist made off with $75 million in drugs from an Eli Lilly warehouse, Dan Burges, director of intelligence at the US division of FreightWatch International, comments on how thieves are shifting their focus from electronics to pharmaceuticals. (*Wall Street Journal*, 17 March 2010)

"I never dreamed that my discovery four decades ago would lead to such a profit-driven public health disaster." Immunologist Richard J. Ablin, University of Arizona College of Medicine, bemoans the $3 billion annual bill for prostate specific antigen screening which, despite a widely reported failure to detect prostate cancer, companies and advocacy groups continue to push. (*New York Times*, 9 March 2010)

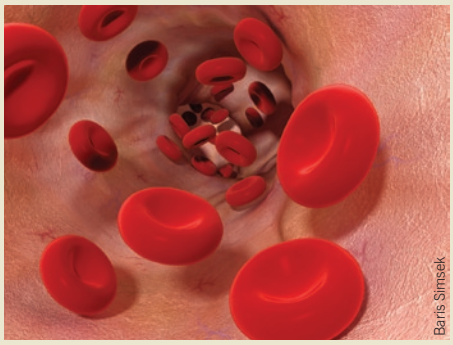
"We are naturally disappointed with the news." Nigel Parker of Ark Therapeutics downplays the European Medicines Agency's request for further costly clinical studies for its glioma gene therapy, days before the company announced it would consider offers. (*The Independent*, 10 March 2010)

"In the future, when all of us have our genomes done, we'll almost certainly have them done in families, because it increases the accuracy of the data." Following publication of a study of Miller's syndrome (*Science* 10.1126/science.1186802, 10 March 2010), Leroy Hood, of the Institute for Systems Biology in Seattle, comments on the power of family sequencing to identify recombination sites, reduce base-call errors and identify rare variations. (*MIT Technology Review*, 11 March 2010)

"What we're trying to prove is that this was a deliberate fraud on their part to block Xenomics, to prevent us from using our own technology." David Jaroslawicz, a lawyer representing TrovaGene (formerly Xenomics), which had anticipated royalties following licensing its Trisomy 21 test to Sequenom, but has been frustrated by delays in commercial launch owing to "employee mishandling of R&D test data and results" at the San Diego–based biotech. (*Xconomy*, 17 March 2010)

Erythropoietins locked into risk management program

The US Food and Drug Administration (FDA) and the makers of erythropoietin stimulating agents (ESAs) have agreed on a formal strategy to reduce the risks associated with these drugs. Starting in March, drug makers will operate under a risk evaluation and mitigation strategy (REMS) requiring healthcare providers who prescribe the drugs for cancer patients to register with the drug makers and enroll in a training program on their use. Applying REMS could help ensure these agents, which have been under scrutiny for their potential to cause tumor growth and shorten overall survival, can remain on the market.



The first erythropoiesis-stimulating agent (ESAs) approved in 1989 revolutionized the treatment of anemia, but over the years safety concerns have been mounting.

Safety concerns for ESAs have been brewing since the FDA convened a session of the Oncology Drugs Advisory Committee (ODAC) in 2004. In December 2006, the company disclosed clinical trials results showing the potential risks of ESAs when used to treat anemia in cancer patients so they won't need blood transfusions. Another ODAC meeting held in 2007 (*Nat. Biotechnol.* **25**, 607–608, 2007) prompted restrictions in coverage by the Centers for Medicare and Medicaid Services, and later a boxed warning on all versions of epoetin alfa, including Amgen's Aranesp and Epogen, and Procrit, sold by Centocor Ortho Biotech of Bridgewater, New Jersey, a subsidiary of Johnson & Johnson. Along the way, the drug makers issued "Dear Doctor" warning letters describing the safety concerns and the appropriate use of the drugs.

Now the REMS program mandates distribution of a medication guide to patients and forces drug prescribers and dispensers to enroll in a special program, called APPRISE, and also document that they have discussed the risks of using ESAs before the start of therapy. Although it's taken three years for FDA and the companies involved to devise a risk mitigation plan, a REMS was not an option when the problems with ESAs became apparent. REMS was authorized under the Food and Drug Administration Amendments Act signed in September 2007 (*Nat. Biotechnol.* **25**, 1189–1190, 2007). "You then had almost six months before [the legislation] actually took effect, so it was already March 2008 before you could even begin a conversation about how to apply a REMS to a product," says Michael McCaughan of the Washington, DC, policy consultancy Prevision Policy. "The issues with ESAs came to a head when the FDA's entire drug safety system was under fire on all sides," he adds. "Then you had the situation where basically every study that was done to look at the issue made the problem worse." McCaughan points out that the ESA REMS is "essentially a manufacturer-FDA agreement" and was not developed in as public a process, involving more interested parties, as it could have been. "It's going to be interesting to see what happens over the course of the next year, when a provider actually has to start using this REMS and has to sign up and register in order to continue to use the products," he says.

In an e-mail, an Amgen spokesperson wrote that "multiple parties have been involved in the development of the ESA REMS, and various healthcare providers were consulted to obtain input at various times. However, the main parties involved in the ESA REMS were the FDA, Amgen and Centocor Ortho Biotech." The not-for-profit healthcare provider Kaiser Permanente, of Oakland, California, submitted a citizen's petition last December requesting that FDA obtain input from providers and others, as required by law. The petition claimed that with the exception of the REMS issued for the use of extended release opioids, the agency has not sought such input. Arnold Friede, an attorney in New York and a former FDA associate chief counsel, has a different complaint with the ESA REMS. Through it, he says, the FDA is in effect "controlling the practice of medicine," which is outside the scope of its authority. "I would submit to you that if an oncologist enrolls in the APPRISE program and signs the enrollment form but then prescribes the drug in a manner that differs from the specific terms of the enrollment form, that the doctor runs a serious risk of negligence *per se* liability should anything go wrong in the treatment," he says.

Mark Ratner *Cambridge, Massachusetts*

IN brief

Stem cells to order

The UK Stem Cell Bank (UKSCB) is relocating to a new building, a move that should boost its growing partnership with the private sector. The Potters Bar–based facility keeps quality-controlled, standardized stocks of stem cell lines that it ships to accredited researchers worldwide together with advice on how to use them. Since its inception in 2004, UKSCB, part of the National Institute for Biological Standards and Control, has dealt primarily with academia, but corporate liaisons are on the rise. One recent client is Stem Cells for Safer Medicines (SC4SM), a public-private collaboration between various UK public stakeholders and private investors including GlaxoSmithKline, AstraZeneca and Roche of Basel. As a not-for-profit company, SC4SM is developing improved human cell toxicology assays to test candidate drugs. Frank Bonner, SC4SM's CEO, says the UKSCB is fulfilling "an absolutely vital role" in ensuring the continuity and quality of their research. The facility now carries about 70 human embryonic lines, with induced pluripotent stem cells under evaluation. According to UKSCB director Glyn Stacey, the Bank serves a number of other companies, mostly for laboratory-based *in vitro* research and toxicology assays, but there are plans to expand the facility's clinical-grade capacity. Currently, neither academic nor corporate users are charged more than delivery, but Stacey says this policy is under review. *Jennifer Rohn*

ReNeuron first in stroke

ReNeuron will be treating the first stroke patients with stem cells later this year in the UK after overcoming a string of regulatory holdups abroad. In February, the Surrey-based company received the go-ahead to start a phase 1 trial from the UK's Gene Therapy Advisory Committee (GTAC) for ReN001, a genetically engineered neural stem cell line originally derived from fetal brain tissue. ReNeuron was the first European stem cell company to go public in 2005, but at the time decided to apply for approval with the US Food and Drug Administration (FDA). The FDA, however, repeatedly delayed approval, prompting the company to apply in 2008 to the UK's Medicines and Healthcare Products Regulatory Agency and subsequently GTAC. CSO John Sinden points out that, "We decided temporarily to discontinue discussions with the FDA since we could not afford to run two phase 1 trials in stroke." The recently approved open label dose escalation safety trial will take place at Glasgow's Southern General Hospital in 12 patients and will also evaluate potential efficacy biomarkers using structural and functional MRI. Chris Mason, who heads the Regenerative Medicine Bioprocessing Unit at University College London, says, "Given the degree of scrutiny by the regulators, the phase 1 study should not be a challenge." Mason adds that any unwelcome surprises are more likely to crop up at later stages, due to the heterogeneity of this patient group. *Susan Aldridge*

EPA releases land-use rule for biofuels to mixed reception

The US Environmental Protection Agency (EPA) has issued long-awaited regulations governing renewable fuel standards, but the reception in political circles has been frosty. Part of the backlash, from both Republican and Democratic politicians, relates to bipartisan opposition to US climate change legislation of any kind. "We need to stop the EPA in its tracks on this and prevent them from simply imposing these over-reaching regulations on all of us," says Rep. Collin Peterson of Minnesota, a Democrat, who is joining forces with Republican members in the House to formally amend the Clean Air Act and block agency actions on regulating greenhouse gas emissions.

What's more, reaction from environmental and consumer groups has been mixed at best. "A few years ago, we saw a large 'tent' coming together for biofuels, but it's broken apart on the land-use issue," says Jim Kleinschmit of the Institute for Agriculture and Trade Policy in Minneapolis. Thus, he says, opposition comes not so much along political party lines as from a "rural-urban" divide, as appears to be the case with Rep. Peterson and his Republican allies. Much of the resistance is attributable to "huge disagreements" about whether or how to include terms dealing with indirect land use when computing the impact of biofuels on greenhouse gas emissions.

"EPA was forced to deal with this [land-use] concept, which has no precedent body of science," says Geoff Cooper, who is vice president of research at the Renewable Fuels Association in Washington. It "cobbled together" models for its draft rule in 2009 and improved those measures for the final rule this February. But the agency "still has significant work to do," he says. "We still feel the land-use [component] is not ready to bear the weight of regulation." Nonetheless, he adds, it is at least "workable for our industry."

Others working in the biofuel sector have given the rules a more positive endorsement (**Table 1**). "We're delighted EPA completed the rule because it sends a message about renewable fuels and the significant percentage [they make up] of liquid transportation fuels," says Michael McAdams, president of the Washington-based Advanced Biofuels Association. The association represents several dozen companies focused on technologies moving beyond first-generation biofuels, mainly corn-based ethanol and biodiesel.

McAdams plays down the importance of the land-use provisions in the EPA rule. "For my group of companies, energy density is a 'must have' and by far the most important feature," he says. "On balance, the EPA rule helps move the ball forward and is more good than bad." Nathanael Greene of the Natural Resources Defense Council (NRDC) in New York is also upbeat: "From the NRDC perspective, the EPA rule is very solid, the agency followed a transparent process, and it is a critical part of the equation," he says. "We need to take this tool and use it to direct our biofuels policy to the least risky biofuels out there." He calls the biofuels life-cycle calculation "very complicated," emphasizing that "not all biofuels are created equal" and noting that some are expected to have a much lower impact in terms of greenhouse gas emissions than others.

Jeffrey L Fox *Washington, DC*

AP Photo/Rich Pedroncelli

Consumer groups have been vociferous in condemning biofuels; here a 10-year old joins his mother in a demonstration in April 2009. The latest EPA regulations have done little to assuage opposition.

Table 1 Recent federal-level biofuel developments

Date	Highlight
February 2010	EPA releases life-cycle analysis of greenhouse calculations for corn ethanol and other biofuels
February 2010	EPA finalizes a rule to implement renewable fuels standards of 36 billion gallons by 2022
March 2010	US Energy Information Administration within the Department of Energy projects ethanol production, which averaged 700,000 bbl/d in 2009, to increase to an average of 800,000 bbl/d in 2010 and 850,000 bbl/d in 2011

Source: US Environmental Protection Agency (http://www.epa.gov/otaq/renewablefuels/420f10007.htm)

New product approvals

Vpriv (velaglucerase alfa)	Shire Pharmaceuticals Group (Basingstoke, UK)	The US Food and Drug Administration approved the new drug application for Vpriv, a gene-activated human glucocerebrosidase, as a long-term enzyme replacement therapy for Gaucher's disease. The enzyme is produced in a human cell line using a proprietary gene-activation technology and has an amino acid sequence identical to the human enzyme.

Shell's billions to convert Brazilian biomass into fuel

Oil giant Royal Dutch Shell in February announced a $12 billion joint venture with Brazilian sugarcane-to-ethanol producer Cosan. In a deal that could lead to large-scale production of advanced biofuels, Shell will contribute its Brazilian fuel distribution network and $1.6 billion in cash, and Cosan, headquartered in Sao Paulo, will devote about two billion liters production capacity per year, with plans to scale up.

The deal represents big oil's largest move into ethanol. But what's got technology analysts particularly excited is the potential for Shell to apply next-generation biofuel technologies to Cosan's production capabilities. Shell says it will contribute to the venture its equity interests in two advanced biofuel developers: Codexis and Iogen, in which the oil giant has 14.7% and 50% stakes, respectively. "This move is a fully integrated play for Shell," says David Berry, a partner with Flagship Venture in Cambridge, Massachusetts.

Codexis, based in Redwood City, California, is developing enzyme products to use as biocatalysts to convert biomass into fuels. The company uses systems biology and gene-shuffling techniques to direct organisms to produce enzymes, such as cellulases, with the desired catalytic activity. Ottawa, Ontario, Canada–based Iogen is developing a cellulosic biomass-to-ethanol conversion process that combines thermal, chemical and biochemical techniques. The companies' technologies enable a wider range of biomass to be converted into fuel, which can create efficiencies when applied to cheap feedstocks and large production processes like Cosan's. "The processing issue is close to being solved by companies like Codexis, so you want access to places with biomass," says Mark Bünger, a biofuels analyst at Lux Research in San Francisco. "There are not many places in the world where feedstocks are cheaper than they are in Brazil."

Shell and Cosan say the deal announced in February is a nonbinding memorandum of understanding and that they intend to negotiate a binding agreement after completing due diligence and regulatory approvals. "It's a good deal for Cosan," says Marco A.P. Lima, director of the Brazilian Bioethanol Science and Technology Center in Campinas. "And I think it is good for Brazil."

Emily Waltz *Nashville, Tennessee*

AP Photo/Lawrence Jackson

Brazil's 30-year-old ethanol fuel program is based on cheap-to-cultivate sugarcane. The new deal for advanced biofuels is for technologies applied to cellulosic biomass.

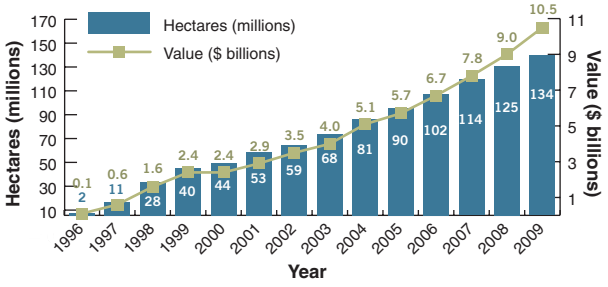
2nd-generation GM traits progress

Andrew Marshall

With 5.6 million new hectares (35%) of transgenic crops, Brazil supplanted Argentina to become the 2nd largest cultivator. China's transgenic plantings shrank, although biosafety certificates were issued for *Bacillus thuringiensis* (*Bt*) rice and phytase maize, clearing the way to commercialization. The first transgenic high oleic soybean was approved in the US, as were disease-resistant varieties of plum and papaya. 2010 plantings of glyphosate-resistant sugarbeet await a US Federal Court ruling.

Historical global area of transgenic crops

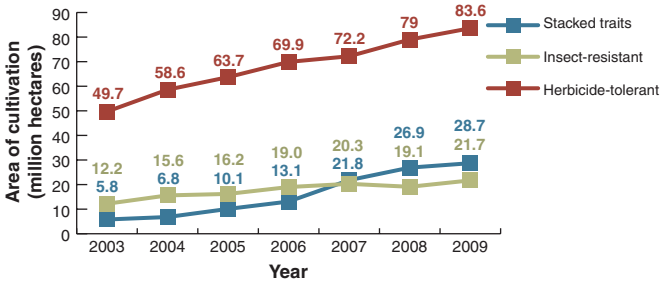
The area planted with transgenic crops rose by 7% in 2009, with their value from seeds/licensing revenues climbing to $10.5 billion.



Source: International Service for the Acquisition of Agri-Biotech Applications, Cropnosis. Value data are explicitly from seeds and licensing revenues rather than from 'crops' themselves.

Global area by transgenic trait

Growth in all categories of transgenic crop continued last year.



Source: International Service for the Acquisition of Agri-Biotech Applications

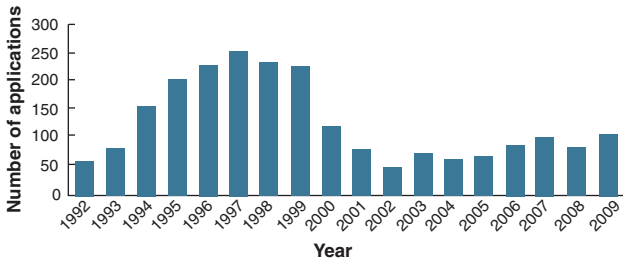
2009 transgenic crop approvals in US and EU

Country	Company/ Institution	Description	Approval type
US	Syngenta Seeds	SYN-IR67B-1/Cotton resistant to lepidopteran pests via expression of Bt cry1Ab	Food
US	Pioneer Hi-bred International	DP-Ø9814Ø-6/Corn tolerant to glyphosate and ALS-inhibiting herbicides via expression of glyphosate acetyltransferase and maize acetolactate synthase	Food/feed
US	Pioneer Hi-bred International	DP-305423/Soybean with high oleic acid, low linolenic acid content via expression of soybean microsomal omega-6 desaturase	Food/feed
US	United States Department of Agriculture	ARS-PLMC5-6(C5)/Plum resistant to plum pox virus via expression of viral coat protein	Food/feed
US	University of Florida	UFL-X17CP-6 (X17-2)/Papaya resistant to ringspot virus via expression of viral coat protein	Environment
EU	Monsanto	MON88017/Corn resistant to lepidopteran pests and glyphosate via expression of Bt Cry3Bb and 15-enolpyruvyl shikimate-3-phosphate synthase	Food/feed
EU	Monsanto	MON89034/Corn resistant to lepidopteran pests via expression of Bt Cry1A.105 and Cry2Ab2	Food/feed
EU	Bayer CropScience	T45/Argentine canola resistant to phosphinothricin herbicide (glufosinate ammonium) via expression of phosphinothricin-N-acetyltransferase	Food/feed

Source: agbios.com

EU transgenic crop field trials

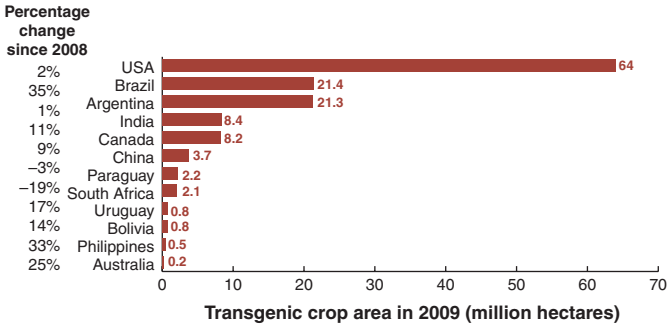
GM field trials in Spain, Czech Republic, Portugal, Romania, Poland and Slovakia increased to 103, with Spain and Romania particularly active.



Source: European Commission Joint Research Center; http://gmoinfo.jrc.ec.europa.eu/gmp_browse.aspx

Global area of biotech crops by country

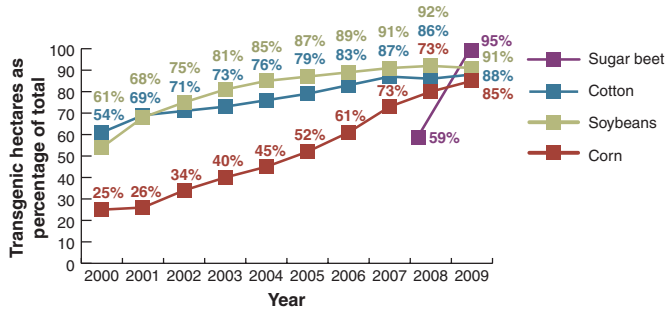
Although transgenic hectarage in China shrank, Brazil continued its rapid growth, supplanting Argentina as the 2nd biggest cultivator of GM crops, with India, Canada, South Africa and Paraguay also expanding cultivation.



Source: International Service for the Acquisition of Agri-Biotech Applications

Transgenic crops as a share of total US crops

Herbicide-tolerant transgenic sugarbeet was rapidly adopted.



Source: National Agricultural Statistics Service; Sugarbeet Growers Association

Andrew Marshall is Editor, Nature Biotechnology

Fresh from the biologic pipeline—2009

Human antibodies take center stage, as they pile up in the list of approved biologics in 2009. Cormac Sheridan reports.

Last year was a vintage year for the biopharmaceutical industry in two respects. Four monoclonal antibodies (mAbs) gained approval in 2009, representing the highest annual number in over a decade. Indeed, this tally equals the record, which was attained in 1998 for several therapeutics that went on to become blockbusters, including Genentech's Herceptin (trastuzumab, for breast cancer) and Centacor's Remicade (infliximab, for autoimmune diseases)[1]. From a technological perspective, it was Medarex that stole the show last year, although it is now part of Bristol-Myers Squibb (BMS), which splashed out $2.4 billion in cash for the Princeton, New Jersey–based biotech. Medarex's UltiMab mAb platform, based on the expression of fully human immunoglobulin sequences in transgenic mice, underpinned all four mAb approvals logged during 2009.

"That is definitely unprecedented," says Janice Reichert, editor-in-chief of *mAbs,* a new journal devoted to antibody therapies, and senior research fellow at the Tufts Center for the Study of Drug Development, in Boston. "To have four approvals in the same year—exclusive of their types—is truly unusual." In addition to these four mAbs, 2009 also witnessed a slew of approvals for H1N1 influenza vaccines in response to the swine flu pandemic as well as the landmark of the first approval for a recombinant protein produced in a transgenic animal (**Box 1**).

A good vintage

The absolute number of new biologic drug approvals in any given year is an imperfect measure of the biopharmaceutical industry's level of innovation. The clinical impact of one highly innovative therapeutic may be of far greater significance than the arrival of several new additions to drug classes that are already well established; in this respect, Johnson & Johnson (J&J; New Brunswick, New Jersey) subsidiary Centocor's (Malvern, Pennsylvania) Simponi (golimumab), approved last April, is the latest in a long line of anti–tumor necrosis factor (TNF-α) inhibitors. Moreover, a single 12-month period can be an arbitrary time frame in which to measure approvals, as unpredictable bureaucratic delays can easily push an expected approval out of one year and into the next. It can thus be argued that, 'spiritually' at least, the interleukin-6 (IL-6) inhibitor Actemra (tocilizumab) belongs to the class of 2009. The drug, developed by Genentech, of S. San Francisco, California, and its parent, Basel-based Roche, gained approval in January 2009 in the European Union for treating rheumatoid arthritis, where it is marketed as RoActemra. Although US approval was expected to follow some time last year, the drug did not get across the finish line until January 11, 2010.

Notwithstanding these caveats, the real value of an annual headcount of new drug approvals lies in the comparison to the totals clocked up in previous years (**Fig. 1**). From that perspective, 2009 was at the upper end of performance in terms of new biologic drugs. The FDA's Center for Drug Evaluation and Research (CDER) approved six novel biologic products in total (excluding H1N1 vaccines; **Box 2, Table 1**), equal to the number achieved in 2003 and just one less than the seven approved in 2002, the best year so far of the new century. However, one of the 2009 total, Dysport (abotulinum toxin A, approved for cervical dystonia), developed by Paris-based Ipsen, was first approved elsewhere as long ago as 1991. Although its molecular weight is much more variable (from 500–900 kDa), it is a similar molecule to Botox (botulinum toxin type A; 900 kDa), which Irvine, California-based Allergan markets as a cosmetic to treat glabellar lines.

The CDER also approved 19 new molecular entities, several of which are distinguished by their mechanism[2], including two new antibiotics—Besivance (besifloxacin; Bausch and Lomb in Rochester, New York) and Vibativ (telavancin; Theravance in S. San Francisco/Astellas in Deerfield, Illinois)—and the second dipeptidyl peptidase 4 (DPP4) inhibitor to reach the market, BMS/AstraZeneca's Onglyza (saxagliptin). The latter figure is in keeping with the trend for the past decade but is well below the annual tallies of small-molecule drug approvals achieved during the mid- and late 1990s. Fifty-three new molecular entities gained approval in 1996, for example[3]. The Center for Biologics Evaluation and Research (CBER) also added several new enzyme replacement therapies, recombinant proteins and new vaccines to the list of new biological products (**Supplementary Table 1** and **Fig. 2**). Across the Atlantic, the European Medicine's Agency (EMEA; London) also approved a sprinkling of novel biologic treatments (**Box 3**).

Human mAbs come of age

The new crop of mAbs represents a resounding level of validation for the HuMab transgenic mouse technology developed by Nils Lonberg, initially at Palo Alto, California–based GenPharm, and then at Medarex, which acquired GenPharm in 1997. The platform was further extended, after a cross-licensing agreement with the Tokyo-based beer company Kirin, which had independently developed a TransChromo mouse strain. UltiMab is based on so-called KM mice, bred from these two lines.

The four 2009 approvals bring to six the total of fully human mAbs on the market.

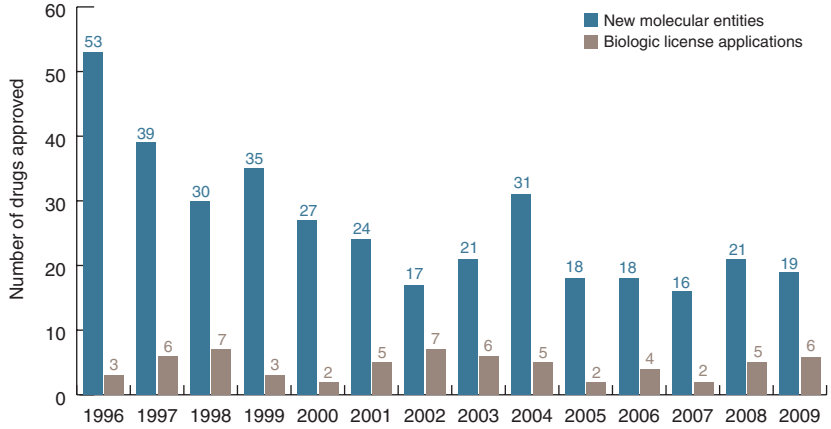


Figure 1 FDA approvals 1996–2009.

Box 1 Waiting in the wings

Three years after gaining approval in Europe for its recombinant version of human antithrombin, an anti-coagulant secreted in the milk of transgenic goats (*Capra aegagrus hircus*), GTC Biotherapeutics of Framingham, Massachusetts, finally secured US approval for Atryn as well, from CBER. The drug approval—and an associated approval by the FDA's Center for Veterinary Medicine of the recombinant DNA construct present in the animals—followed the publication in January 2009 of the agency's final guidance on the regulation of transgenic animals in commercial use. Atryn, which inhibits the activity of the coagulation proteins Thrombin, factor Xa and factor IXa, is aimed at patients with hereditary antithrombin deficiency, a rare autosomal dominant disorder, which can lead to the formation of life-threatening blood clots. Although normal blood thinners can be used for routine therapy, antithrombin is indicated for patients undergoing surgery or before, during or after childbirth.

Although now commonplace in biomedical research, the use of transgenic animals for 'pharming' or protein production has, so far, remained a fringe area. The world's first transgenic mammal, a bull called Herman, genetically engineered to produce offspring expressing the human protein lactoferrin, was born in The Netherlands two decades ago. Although he died in 2004, the product is still undergoing development—as a nutraceutical. Herman's owner, Leiden-based Pharming—whose roots lie in the European operations of GenPharm International—is still seeking regulatory approval for Rhucin, a C1 esterase inhibitor produced in the milk of transgenic rabbits, in development for hereditary angioedema patients.

GTC is positioning its platform as a highly scalable, "game-changing" technology, with the potential to produce 'biosimilar' and 'biobetter' antibodies in large quantities. For now, however, adoption of transgenic animals as drug bioreactors remains limited, according to Janice Reichert, of the Tufts Center for the Study of Drug Development. "There are not many people doing work in the transgenic animal area," she says. *CS*

Before 2009, Amgen's Vectibix (panitumumab), an inhibitor of epidermal growth factor receptor approved for colon cancer, and Humira (adalimumab), a TNF-α inhibitor marketed by Abbott Laboratories, of Abbott Park, Illinois, for autoimmune disorders, were the sole representatives of this category. Vectibix was generated by means of the XenoMouse technology developed by Abgenix, of Fremont, California, and now owned by Amgen, of Thousand Oaks, California. Humira's origins lie in the phage display technology developed by Cambridge Antibody Technology, of Cambridge, UK, and now owned by AstraZeneca, of London.

And there's more to come. "The four drugs approved in 2009 really represent only the tip of the iceberg for our transgenic mouse platform. There are a lot of exciting drugs behind these in clinical development, and we continue to use the platform for drug discovery," says Lonberg, now senior vice president, biologics discovery at BMS. UltiMab's new owner is continuing to invest in further development of the platform. "We continue to make new strains of engineered mice that add to the basic platform, but, just as important, we also continue to invest in the infrastructure necessary to characterize and identify lead drugs derived from the mice. And we have developed our own proprietary antibody-drug conjugate technology that extends the range of applications for human antibody–based drugs," Lonberg says. For example, MDX-1203, an anti-CD70 antibody conjugated via a peptide-based linker to a prodrug of CC-1065 (rachelmycin), entered phase 1 testing for advanced/recurrent renal cell carcinoma and relapsed/refractory B-cell non-Hodgkin's lymphoma last July.

Copenhagen-based Genmab is another beneficiary of UltiMab's recent coming of age. Formed in 1999 as a European spin-out from Medarex, Genmab also holds rights to the UltiMab platform, and it can lay direct claim to one of last year's four antibody approvals, Arzerra (ofatumumab). It licensed the compound, an anti-CD20 mAb, to London-based GlaxoSmithKline in late 2006. The entire development process, from generating a hybridoma cell line to final approval, took seven years, seven months and five days, says Jan van de Winkel, Genmab's chief scientific officer. "That's a record time."

The drug, which gained approval for treating patients with chronic lymphocytic leukemia (CLL), who have failed other therapies, is a competitor to an older, chimeric anti-CD20 antibody, Rituxan (rituximab) from Genentech and Cambridge, Massachusetts–based Biogen Idec, which also recently gained marketing approval for CLL. Arzerra, though, binds a different epitope on the B-cell CD20 antigen, which results in more efficient activation of immune effector functions, such as antibody-dependent cellular cytotoxicity and complement activation, according to van de Winkel. "That's very distinct from all other CD20 antibodies," he says.

Genmab offers an alternative route to UltiMab-generated antibodies for companies that may, for competitive reasons, be unable or unwilling to partner with BMS. Genmab is not entitled to license the technology, but it can out-license antibodies it generates using the technology, unless Medarex has entered an exclusive agreement with a third party on a particular target—a practice it has not generally adopted up to now. "Of course,

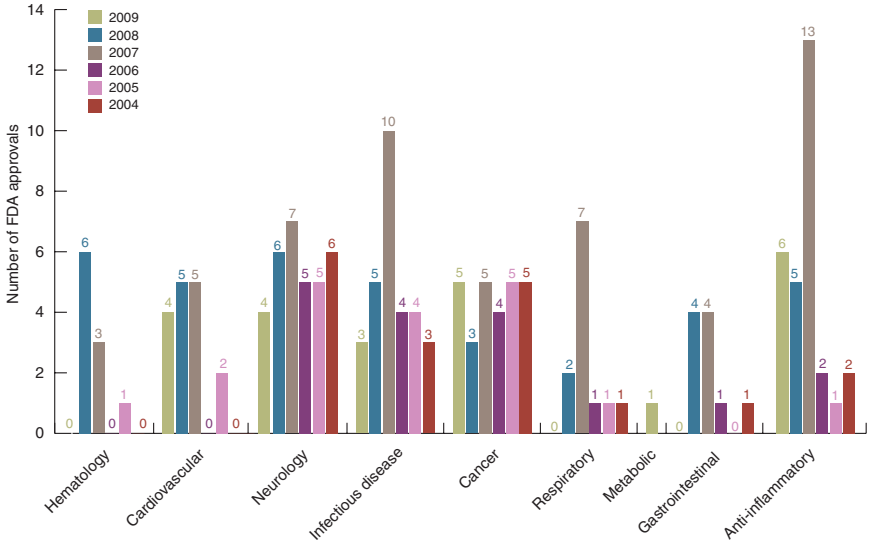


Figure 2 FDA new molecular entities and biologics license applications since 2004 according to therapeutic indication. Tallied numbers may be in more than one indication. Source: Biocentury, BCIQ.

BMS can make a corporate decision to make the platform less available, but that should not apply to Genmab," van de Winkel says. "Contractually, we have access to the platform."

That access—coupled with Genmab's focus on drug discovery and development instead of sales and marketing—should ensure that the UltiMab technology becomes more widely disseminated than the XenoMouse platform, over which Amgen has kept a tight rein since its $2.2 billion acquisition of Abgenix. A second XenoMouse-derived human mAb, Prolia (denosumab), which Amgen also plans to commercialize, could join Vectibix on the market very shortly. The FDA is due to complete its review of an application for the drug's approval as an osteoporosis treatment by July 25. But there are no XenoMouse-derived antibodies in late-stage development at other firms.

To market, to market

All of the antibodies approved in 2009 are being commercialized by large pharmaceutical companies. Ilaris (canakinumab), an interleukin-1β inhibitor, was fully developed in-house at the Boston headquarters of the Novartis Institutes of Biomedical Research (NIBR). Ilaris is not a first-in-class drug, as two other IL-1β inhibitors already exist. Kineret (anakinra), a recombinant version of the naturally occurring protein IL-1 receptor antagonist, was originally developed by Amgen and is now marketed by Stockholm-based Orphan Biovitrum. The fusion protein Arcalyst (rilonacept), an IL-1 trap that acts as a soluble decoy receptor, is also marketed by Regeneron Pharmaceuticals, of Tarrytown, New York.

Ilaris is, however, the first antibody to be successfully developed against the target. "It is very specific and has a very long half-life and can suppress IL-1 for many months," says NIBR president Mark Fishman. "That's what we know is different from the other two." Its development path is emblematic of the drug development philosophy Fishman espouses, which is based on gaining rapid approval in a rare condition before gaining further approvals in broader indications that have a similar underlying molecular mechanism.

Ilaris has initially gained approval in familial cold auto-inflammatory syndrome and in Muckle-Wells syndrome, two of the three disorders of the innate immune system collectively known as cryopyrin-associated periodic syndrome. The condition is caused by an autosomal dominant mutation in *NLRP3*, which encodes cryopyrin, an activator of the inflammasome, a multiprotein complex involved in regulating the inflammatory response. The mutation results in a damaging overproduction of IL-1β. The antibody also has potential in systemic juvenile idiopathic arthritis, and more common disorders, such as some forms of gout, chronic obstructive pulmonary disorder and type 2 diabetes. But Fishman sees value—commercial as well as clinical—in addressing rare disorders in any case. "I believe at the end of the day, it's almost inevitable you will make a profit on a good medicine."

J&J's Centocor Ortho Biotech arm can claim another two of the six biologic approvals, one of which, Stelara (ustekinumab), is a first-in-class drug. It inhibits the pro-inflammatory cytokines IL-12 and IL-23 by binding the p40 subunit common to each. That prevents their binding to the cell-surface receptor chain, IL-12 β1, and triggers their respective immunological cascades. The two cascades are ultimately distinct, as the IL-12 p35 and the IL-23 p19 subunits recognize different receptors and activate alternative inflammatory pathways. IL-12 is associated with a T-helper (Th) type 1 response (Th1), whereas IL-23, which has variously been described as the master switch in both psoriasis and Crohn's, is associated with a Th17 response. Stelara has gained approval initially as a treatment for plaque psoriasis, but studies in Crohn's disease and psoriatic arthritis are also underway. However, the molecule failed to demonstrate efficacy in multiple sclerosis[4]. Its long-term safety profile is not well understood, either. The FDA approval required J&J to put in place a Risk Evaluation and Mitigation Strategy to evaluate and mitigate the risk of serious infection, malignancy and the development of reversible posterior leukoencephalopathy syndrome, a neurological condition associated with taking chemotherapy and transplant rejection drugs.

Box 2 Pandemic vaccines and beyond

Vaccine R&D at multinationals, such as Paris-based Sanofi-aventis, GlaxoSmithKline (GSK), Merck and Wyeth, of Madison, New Jersey, continues to experience something of a renaissance, with global sales reaching $38 billion last year. With the emergence of the threat of an H1N1 influenza pandemic, several companies obtained approvals for egg-cultured, inactivated vaccines, including both a monovalent and trivalent vaccine (Agriflu) from Novartis, and monovalent vaccines from GSK subsidiary ID Biomedical, Sanofi-aventis and CSL Behring. In September, AstraZeneca's subsidiary MedImmune gained approval for a live H1N1 Flumist vaccine. Further afield, Beijing-based Sinovac Biotech also received approval for an H1N1 vaccine from the Chinese authorities.

A number of companies also successfully launched vaccines against other pathogens. GSK's Hiberix, a *Haemophilus influenzae* type b conjugate vaccine was approved for meningitis last August. The same company's Cervarix—a vaccine comprising human papilloma virus types 16 and 18 in L1 virus-like particles that has been long touted as a competitor for Merck's Gardasil vaccine—finally received FDA approval in October. Seven months earlier, Vienna-based Intercell Biomedical's Ixiaro (Japanese encephalitis SA14-4-2 virus vaccine) was approved for the prevention of Japanese encephalitis disease.

Andrew Marshall

Table 1 Recombinant biologic approvals in 2009[a]

Company/partner	Product	Indication
Centocor Ortho Biotech	Simponi; human mAb specific for TNF-α	Rheumatoid arthritis, psoriatic arthritis and ankylosing spondylitis
Genzyme	Atryn (antithrombin III); recombinant antithrombin III, produced in transgenic goats	Blood clots in patients with antithrombin deficiency
Fresenius Biotech and TRION Pharma	Removab; rat/murine bispecific mAb with one arm specific to EpCAM and the other to CD3	Approved in Europe for malignant ascites in patients with EpCAM-positive carcinomas
Novartis	Ilaris; human mAb specific for IL-1β	Cryopyrin-associated periodic syndromes
Centocor Ortho Biotech	Stelara; human mAb specific for the p40 subunit of IL-12	Moderate-to-severe plaque psoriasis
GSK	Arzerra; human mAb specific for CD20	CLL
Dyax	Kalbitor; plasma kallikrein inhibitor	Hereditary angioedema

[a]For a more complete list, see **Supplementary Table 1**.

Box 3 Biologics at the EMEA

Although many companies have traditionally targeted the US market for approval of novel products, others have chosen the European Medicine Agency (EMEA) in London as the starting point. For example, Fresenius Biotech of Homburg, Germany and Munich-based TRION Pharma obtained approval for Removab (catumaxomab) from the EMEA last April for the treatment of malignant ascites in individuals with epithelial cell adhesion molecule (Ep-CAM)-positive carcinomas. Removab is a product of TRION's bispecific mAb program. It is rat/murine bispecific mAb with one arm specific to EpCAM and the other to CD3. As well as bringing $CD3^+$ T-cells into proximity with $EpCAM^+$ epithelial tumor cells, it is also thought to function via the interaction of its Fc fragment with FcγRI/III on accessory immune cells. Last June, the EMEA also approved a cell therapy from Leuven, Belgium–based TiGenix for cartilage defects of the femoral condoyle of the knee. The therapy, ChondroCelect, is prepared by extracting chondrocytes from a healthy region of a patient's cartilage; these are then expanded *ex vivo* and then re-implanted surgically. *AM*

"We have already compiled three years of safety data from our ongoing five-year psoriasis studies, and these results have demonstrated a high level of efficacy in patients with moderate to severe psoriasis with a favorable benefit-risk profile," says Centocor Ortho Biotech spokesman Brian Kenney. "Additional data from these five-year trials, as well as data from other ongoing trials in psoriasis, psoriatic arthritis, Crohn's disease and sarcoidosis, will aid in establishing the long-term safety profile of Stelara."

Simponi, J&J's second biologic to gain approval in 2009, is the fifth TNF-α inhibitor to reach what is now a lucrative but crowded market. It is going up against several mature product franchises, including Amgen's Enbrel (etanercept), Abbott's Humira and Remicade, marketed in the US by J&J and in Europe by Merck of Whitehouse Station, New Jersey, following its acquisition of Kenilworth, New Jersey–based Schering-Plough. It is also pitted against a more recent newcomer: Cimzia (certolizumab pegol), a pegylated Fab' antibody fragment, marketed by UCB, of Brussels. Stelara, which is administered by subcutaneous injection, can, like Cimzia, be self-administered by patients. It also offers a more convenient dosing schedule than other TNF-α inhibitors, as only five injections per year are necessary. However, it is not expected to offer significant clinical advantages over its longer-established competitors[5].

More to come

Apart from Ipsen's Dysport, the only other recombinant protein to gain CDER approval in 2009 was Kalbitor (ecallantide), developed by Cambridge, Massachusetts-based Dyax for treating acute attacks of hereditary angioedema, a condition caused by a deficiency in C1 esterase inhibitor, which normally dampens the complement cascade. CSL Behring, of King of Prussia, Pennsylvania, gained CBER approval in the same indication, for Berinert, a plasma-derived version of C1 esterase inhibitor.

In mainstream indications, however, the dominance of mAbs on the list of biologic approvals is set to continue. "The rate of entry [of antibodies] into clinical studies has gone up very dramatically since the late 1990s," says Reichert. Back then, less than 20 were entering the clinic annually, whereas around 40 antibodies are doing so at present. The equivalent figures for other types of recombinant proteins have remained "pretty flat" in contrast, she says. The current clinical pipeline of biologic drugs includes 240 mAbs and 120 proteins in various other categories, including alternative scaffolds, fusion proteins and recombinant versions of naturally occurring plasma proteins. Based on previous success rates of ~20%—and on typical development timelines of seven to eight years—she estimates that an additional 48 mAbs could reach the market within that time frame. "In seven or eight years, will we have 48 more? That's an open question. If it happens, that would be pretty good."

Cormac Sheridan, Dublin

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

1. http://www.landesbioscience.com/journals/mabs/about#background
2. Anonymous. CDER New Molecular Entity (NME)/ New BLA Calendar Year Approvals As of December 31, 2009 (FDA CDER, Washington, DC, 2009). <http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/DrugandBiologicApprovalReports/NMEDrugandNewBiologicApprovals/UCM091096.pdf>
3. Hughes, B. *Nat. Rev. Drug Disc.* **9**, 89–92 (2010).
4. Segal, B.M. *et al. Lancet Neurol.* **7**, 796–804 (2008).
5. Pappas, D.A. *et al. Nat. Rev. Drug Disc.* **8**, 695–696 (2009).

Biotech in a blink

The eye, and particularly the retina, has become a favored testing ground for new biologic drugs. How well novel nucleic acid and cellular therapies work in retinal disease could determine their expansion to other indications. Ken Garber reports.

Long neglected by the drug industry, the eye is now a favored target for new drugs, especially novel biological therapies. The stunning success five years ago of anti-angiogenic therapy in the neovascular (wet) form of age-related macular degeneration (AMD) has accelerated drug development for retinal pathologies, which each year condemn millions to blindness.

As a largely self-contained organ that offers easy access and partial immune privilege, the eye is prized as a testing site for novel biologics. At the same time, research is demystifying AMD and other retinal diseases. "There's been a revolution in retinopharmacology," says Anthony Adamis, vice president and global head of ophthalmology at Genentech in S. San Francisco, California. "The science has matured."

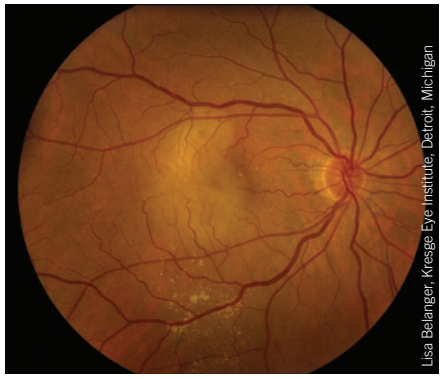
Folkman's other legacy

Anti-angiogenesis therapy has encountered far greater success in retinal disease than in cancer. In 1948, ophthalmologist Isaac Michaelson hypothesized that a soluble 'factor X' produced in the retina was the main cause of the new blood vessel growth, a hallmark of retinal degeneration in diabetes. In the early 1960s surgeon Judah Folkman originated the angiogenesis concept in cancer and postulated that tumors secrete something that recruits new blood vessels. Folkman called it TAF, for tumor angiogenesis factor. He speculated almost 40 years ago that TAF and factor X were the same thing.

It proved to be vascular endothelial growth factor (VEGF). Discovered twice in the 1980s—first as 'vascular permeability factor' and then separately as a stimulator of angiogenesis—VEGF soon became the main factor X suspect. In the mid-1990s, work by Adamis in Folkman's laboratory and by others confirmed VEGF as the cause of new blood vessel growth in diabetic eyes. By then, it was clear that TAF was also VEGF.

Researchers at Genentech led by Napoleone Ferrara, who had cloned VEGF, set about developing anti-VEGF antibodies for retinal diseases as well as for cancer. In 2004, the US Food and Drug Administration (FDA) approved Avastin (bevacizumab) for colorectal cancer, but treatment extends median survival less than five months. In the meantime, Genentech brought forward Lucentis (ranibizumab), a Fab fragment of Avastin, hypothesizing that the smaller molecule would better penetrate the retina than the full-length antibody.

The phase 3 results of Lucentis in AMD, reported in 2005, stunned the ophthalmology community[1]. "It was a field littered with failed therapies," recalls Adamis. "So…even I, who had seen animal data, was surprised by how well it worked." In the two trials, 34% and 40% of AMD patients on Lucentis improved their vision by three lines or more on an eye chart, compared with 5% and 6% for sham injections and photodynamic therapy, respectively. Almost all Lucentis patients had vision improvement or stabilization—some even recovered 20/20 vision—whereas most on the control arms lost vision. "The difference is as dramatic as almost anything one sees in any medical or scientific publication," says Robert Frank, an ophthalmologist at Wayne State University in Detroit.



Retina showing age-related macular degeneration (AMD). (Macula is bright area in center.) AMD is the leading cause of vision loss in the developed world.

The FDA approved Lucentis in June 2006. By then, ophthalmologists *en masse* were already using Avastin off-label to treat AMD, with the help of compounding pharmacies, which can safely divide up a standard vial of Avastin into small doses for eye injection. A Lucentis injection costs >$2,000, compared with <$50 for Avastin. Genentech in vain has tried to discourage Avastin use in the eye. According to data presented at the fall 2009 Retina Congress in New York, 58% of Medicare fee-for-service AMD patients get Avastin, and only 42% get Lucentis. Reports to date suggest Avastin is just as effective as Lucentis, and an ongoing National Eye Institute–sponsored head-to-head study should provide definitive data.

Either way, VEGF blockade dominates wet AMD treatment. Lucentis earns Genentech >$2 billion a year, a franchise destined to grow, as roughly one in three people will be affected to some degree by AMD by age 75. A similar story is now playing out in diabetic macular edema (DME), present in 9% of people with diabetes over the age of 45. With four separate phase 3 trials nearing completion, the community is optimistic that Lucentis is superior to laser photocoagulation. "Everybody thinks that it is likely to be effective [for DME]," says Frank. Ophthalmologists already use Avastin for DME.

There are few cases in modern medicine where a treatment has so quickly and thoroughly transformed a field as anti-VEGF antibodies in retinal disease. It's a startling vindication of Judah Folkman's longstanding faith in the anti-angiogenesis approach.

A crowded pipeline

But anti-VEGF is not the last word in these diseases. Although the success of Lucentis and Avastin has led some companies to shut down their ophthalmology programs, it has given birth to many others. About 40 compounds are in development for retinal diseases (**Table 1**). The market opportunity is there, because Lucentis and Avastin have several problems. One is the need for monthly injections into the eye. This can lead to endophthalmitis, a devastating infection that can destroy the eye. It also risks hemorrhage or retinal detachment. The more injections, the higher the risk.

Less frequent injection is the explicit promise of VEGF Trap (aflibercept), from Regeneron Pharmaceuticals in Tarrytown, New York. VEGF Trap consists of domains from both VEGF receptors 1 and 2 fused to an IgG backbone, and it binds VEGF-A—the VEGF family member most often implicated in angiogenesis—much more strongly than Lucentis, Avastin or VEGF-A's native receptor. And, unlike Lucentis, it also binds VEGF-B and placental growth factor, which appears to play a role in angiogenesis in the eye and inflammation. The dose is four times that of Lucentis. "The fact of a higher dose, a greater potency and [that it affects] other VEGF family members give theoretical advantages to VEGF Trap," says Peter Campochiaro, an ophthalmologist at Johns Hopkins University. Two phase 3 head-to-head trials against Lucentis in wet AMD are nearly complete.

Regeneron hopes that VEGF Trap, because of

its high binding affinity for VEGF, can be given every two months and still match monthly Lucentis in efficacy. If that schedule pans out, VEGF Trap only needs to be as good as Lucentis to seize market advantage. That outcome is not a foregone conclusion. VEGF Trap "is not a slam-dunk, like VEGF [blockade] was," says Michael Tolentino, a researcher at the Center for Retina and Macular Disease in Winter Haven, Florida. But the VEGF Trap phase 3 results are eagerly awaited. A new drug application filing in wet AMD is possible early next year. FDA approval will likely lead to a fierce marketing war between Regeneron and Roche/Genentech.

Another strategy is to induce blood vessel regression. VEGF blockade can halt angiogenesis and plug vessel leakage, but doesn't eliminate those vessels—one reason for the need for constant retreatment. Pericytes, supporting cells that appear around small blood vessels, are recruited late in the angiogenesis process and are dependent on platelet-derived growth factor (PDGF)[2]. Combining a PDGF or PDGF receptor inhibitor with a VEGF inhibitor is now being tried for AMD. Anti-PDGF treatment strips off the pericytes, Tolentino explains, "and you have the endothelial cells that are left over, which are sensitive to anti-VEGF inhibition." Several small molecules and biologicals are in the clinic (**Table 1**).

Extended release devices, ranging from tiny capsules to intraocular gels, could eventually solve the frequent injection problem. They are already used to deliver steroids to the eye. But delivery of proteins is much harder because proteins tend to oxidize and denature over time. "That's the dogma, but progress is being made on that front," says Adamis. Genentech is collaborating with SurModics in Eden Prairie, Minnesota, on a sustained delivery Lucentis formulation, and other companies are pursuing their own technologies.

siRNA: abandon all hype

One great hope for a more durable AMD therapy was RNA interference (RNAi). The 2001 discovery that 21-nucleotide RNA duplexes could specifically silence mammalian genes immediately raised the prospect that siRNAs could treat disease. RNAi companies soon looked to the eye for proof of principle. By 2007, two siRNA molecules, bevasiranib from Opko Health in Miami (which had acquired Acuity Pharmaceuticals) and AGN-745 from Allergan of Irvine, California (licensed from Sirna), were in phase 3 and phase 2, respectively, for wet AMD. Bevasiranib, which targets VEGF, was the first RNAi therapeutic tested in humans, and the first to reach phase 3. AGN-745 targets the VEGF receptor. The explicit hope for bevasiranib was for long-term suppression of VEGF and less frequent injections than current antibody therapy.

Both siRNA trials were terminated last year. In a terse press release in March, Opko reported that an independent data monitoring committee had recommended termination, and in May Allergan quietly returned AGN-745 to Merck, which had bought Sirna in 2006. Neither drug is now in development.

Because clinical results for the drugs have not been published, reasons for their demise are unknown. (Opko did not respond to phone queries, and Allergan says only that its drug was safe but ineffective.) One big issue for siRNA is target tissue penetration, but Opko published data in 2008 showing that bevasiranib, injected into the vitreous of rabbit eyes, reached high concentration in the retina, and Sirna published similar results in mice. Durability of effect is a more likely reason for failure. Allergan's siRNA, in mice, disappeared from the retina after five days. Campochiaro, who did preclinical work on the drug, says "How can this possibly work when you inject it into the eye, the effect only lasts five days?" Yet Allergan took the drug to phase 2. "I saw some of the results of that phase 1 study," says Campochiaro. "Basically there was nothing there."

Campochiaro considers naked siRNA impractical for anything but short-term acute problems. Tolentino, who invented bevasiranib, says the drug's effects should persist via activation of the RNA-induced signaling complex. But he doesn't know why it failed in phase 3.

Eyes on gene therapy

Gene therapy is another experimental approach that, in theory, could provide a long-term solution—perhaps even a cure. The eye is an attractive place to deliver gene therapy vectors, mostly because it's a largely self-contained organ. The blood-retina barrier mostly keeps the drug in and the immune system out, and the transparency of the cornea, lens and intraocular fluid allow close monitoring of effects.

The eye has provided the setting for one of gene therapy's biggest successes. Jean Bennett's group at the University of Pennsylvania in Philadelphia, beginning in 2007, gave a single subretinal shot of an adeno-associated virus vector carrying the gene encoding RPE65 (retinal pigment epithelium 65 kDa protein) to the eyes of 12 people with Leber's congenital amaurosis (LCA). LCA is a rare progressive inherited disorder that leads to blindness by early adulthood. Last year, Bennett's group reported that all patients regained visual function, some dramatically[3].

The spectacular success of the LCA trial, though, will be hard to repeat for AMD. In LCA, mutations in RPE65 disable the "visual cycle," the enzymatic recycling of 11-*cis*-retinal, a photosensitive derivative of vitamin A essential for the conversion of light into neuronal signaling. Photoreceptors, which capture visual information and transmit it to the brain, remain mostly intact, and gene therapy in LCA succeeds by simply restoring the visual cycle. "To try to translate that finding into other eye diseases, and in fact other diseases in general, is a big leap," says Tolentino.

Genzyme is taking on the challenge of AMD gene therapy in a phase 1 trial of soluble Flt1 (VEGF receptor 1) delivered by adeno-associated virus. Expressed by retinal cells, the soluble receptor binds VEGF and inhibits new blood vessel formation in the choroid. The company, as of late February, was screening candidates and expected to soon start dosing patients. The Cambridge, Massachusetts–based company reported efficacy lasting five months in a primate model, and began enrolling patients in March.

Meanwhile, Oxford BioMedica in Oxford, UK, plans to begin clinical trials this year with lentiviral delivery of the angiogenesis inhibitors angiostatin and endostatin in AMD and diabetic retinopathy, in collaboration with Paris-based Sanofi-aventis. A trial for gene replacement in Stargardt's disease, a single-gene disorder (caused by mutations in a gene for transporter of Vitamin A intermediates), is also imminent.

Focusing beyond VEGF

Many companies are now looking beyond VEGF, because VEGF blockade does not arrest the underlying disease. "It appears that once VEGF is upregulated, in most patients it continues, so you have to continue to block it," says Campochiaro.

Another concern is that long-term angiogenesis inhibition may have harmful side effects not only in the eye (retinal survival depends on a constant blood supply) but also elsewhere (e.g., the heart, in ischemic conditions like heart disease) if inhibitors leak into the systemic circulation. VEGF blockade could also interfere with neurotrophic and neuroprotective functions of the growth factor observed in animals[4]. Such side effects have not been reported in humans, but it's still early.

Insights from genetics and molecular biology are clarifying our understanding of why AMD occurs and providing new targets. The general picture is this: Some triggering event leads to the deposition of cellular and acellular debris in Bruch's membrane, which separates the RPE from the choroid, the supplier of blood to the outer layers of the retina. (The RPE in turn blankets the retina's photoreceptors.) Debris deposition leads to the formation of drusen, yellow deposits of protein and lipid materials, along the RPE. In 2000, complement proteins were found in the drusen of donor eyes. "That

Table 1 Selected investigational therapies for retinal disease

Company	Agent	Mechanism	Indication	Stage
Roche/Genentech, Novartis (outside US)	Lucentis	Anti-VEGF Fab fragment	Wet AMD DME	Approved 2006 Phase 3
Eyetech (Palm Beach Gardens, Florida)/ Pfizer	Macugen (pegaptanib)	Anti-VEGF165 aptamer	Wet AMD DME	Approved 2004 Phase 3
Regeneron	VEGF Trap-eye	Anti-VEGF soluble receptor	Wet AMD CRVO	Phase 3
Neurotech (Lincoln, Rhode Island)	NT-503	Encapsulated CTNF	Dry AMD/GA	Phase 2 complete
GlaxoSmithKline (Philadelphia)	GW786034 (pazopanib)	Kinase inhibitor targeting VEGF, PDGFR and c-kit	Wet AMD	Phase 2
Ophthotech (Princeton, New Jersey)	E10030	Anti-PDGF aptamer	Wet AMD	Phase 1
Alexion Pharmaceuticals (Cheshire, Connecticut)	Soliris (eculizumab)	Anti-complement component 5 (C5) monoclonal antibody	Dry AMD/GA	Phase 2
Quark Pharmaceuticals (Fremont, California)/Pfizer	PF-04523655	siRNA targeting TRP801	Wet AMD DME	Phase 2 Phase 2
Potentia (Louisville, Kentucky) Alcon (Fort Worth, Texas)	POT-4	Peptide inhibitor of C3	Wet AMD	Phase 1 complete
Ophthotech	ARC1905	Anti-C5 aptamer	Wet AMD Dry AMD	Phase 1 Phase 1
Genzyme	AAV2-sFLT01	Soluble VEGF receptor gene therapy	AMD	Phase 1
Genentech	anti-factor D	Anti-complement factor D antibody	Dry AMD/GA	IND submitted
Centocor (Johnson & Johnson; Radnor, Pennsylvania)	hUTC	Human umbilical tissue–derived cells	Retinitis pigmentosa	Phase 1 complete
Advanced Cell Technology (Worcester, Massachusetts)	hESC-derived RPE	RPE cell therapy	Stargardt's macular dystrophy	IND submitted
Optherion (New Haven)	rCFH	Recombinant complement factor H	AMD	Preclinical
Oxford Biomedica/Sanofi-aventis	RetinoStat	Endostatin, angiostatin gene therapy	Wet AMD	Preclinical
University College London/Pfizer	hESC-derived RPE	RPE cell therapy	AMD	Preclinical

AAV, adeno-associated virus; AMD, age-related macular degeneration; DME, diabetic macular edema; CRVO, central retinal vein occlusion; CTNF, ciliary neurotrophic factor; hESC, human embryonic stem cell; GA, geographic atrophy; IND, investigational new drug; PDGF, platelet-derived growth factor; RPE, retinal pigment epithelium; VEGF, vascular endothelial growth factor.

was one of the early indicators that maybe there's some sort of inflammatory process that's going on in the genesis of drusen," says Campochiaro.

The drusen damages Bruch's membrane and the RPE so new blood vessels in the choroid can invade the retina. In wet AMD, leakage from these new vessels disrupts vision in the macula, the central portion of the retina responsible for high-resolution visual acuity. Finally, fibrotic scarring and photoreceptor damage lead to vision loss, including blindness. VEFG blockade decreases leakage from the new blood vessels, so that the retina dries out and vision recovers.

In 2005, three groups, using different methods, including the first successful genome-wide association study, strongly implicated complement factor H mutations in AMD. Follow-up studies found three more complement-related AMD genes. Together, the four genes account for most of the genetic risk of the disease.

The overwhelming evidence implicating complement has driven a new race to develop complement inhibitors for AMD. At least two companies have complement inhibitors in clinical trials (**Table 1**), with perhaps a half-dozen others at earlier stages. This is a new field—the first complement inhibitor was approved only in 2007 for a rare blood disorder—and many questions await answers, including which complement protein to target, and how to block it.

The most important question is whether complement inhibition will work at all in AMD. "The scientific data [on AMD causality] is compelling," says Philip Rosenfeld, an ophthalmologist at the University of Miami. "The question is whether inhibiting complement at this late stage will modify progression of disease." Tolentino has his doubts. "It's difficult to stop this pathway, depending on when you get it," says Tolentino. "Have you ever tried to stop an explosion by capturing all the shrapnel?"

Complement inhibitors "will demonstrate some effect," predicts Campochiaro. "It'll be really fortunate if that effect can be on the neovascular form of the disease because it's going to be extremely difficult and time consuming to show that [they alter] the degenerative form of the disease."

Embryonic stem cells on trial

This degenerative form, known as atrophic or dry AMD, comprises ~90% of cases, and there is no treatment, although the disease progresses more slowly than wet AMD. Instead of new blood vessels forming and leaking as in the wet form, in dry AMD the photoreceptors just wither and die for unknown reasons. Dry AMD, in its late stages, produces 'geographic atrophy': patterned patches of retinal degeneration that appear almost exclusively in the macula, and lead to vision loss.

And dry AMD pathology progresses, even in the wet form of the disease. Some ophthalmologists worry that anti-VEGF therapy makes things worse. "Does anti-VEGF therapy actively accelerate the growth of dry AMD? No," says Rosenfeld. "Geographic atrophy appears to progress exactly as you would anticipate based on the natural history of the disease." The bad news, though, is that anti-VEGF treatment is not likely to stop geographic atrophy in the long run—a growing issue as patients go out several years on Lucentis or Avastin. "Vision loss from dry AMD will eventually erode the visual acuity benefits of anti-VEGF therapy over time," says Rosenfeld.

One possible solution is cell therapy. Many research groups—and some companies—are working on systems to transplant stem cells of various types and differentiation states into the human retina to treat AMD. As with gene therapy, the eye is an excellent site for stem cell transplantation, because of its transparency and

partial immunoprivileged status. And relatively few transplanted cells should suffice, compared with, say, the spinal cord, heart or brain. "The cells of the back of the eye, you're talking about a single cell layer," says Pete Coffey, who researches retinal transplantation at University College London. "Tens of thousands rather than tens of millions of cells."

Stem cell therapy is just one hope for the retina, but the retina may be the best hope for stem cell therapy. The eye could determine the fate of the whole field. Last November, Worcester, Massachusetts–based Advanced Cell Technology (ACT) filed an investigational new drug (IND) application for a phase 1/2 trial using human embryonic stem cell (hESC)-derived retinal cells to treat Stargardt's disease, a leading cause of juvenile blindness. The company plans to file for AMD as well. If the Stargardt's disease IND is approved, it could trigger the first hESC human trial. (Geron's IND for hESC derived oligodendrocytes in spinal cord injury is on clinical hold from the FDA, pending animal studies.)

"To go in and try and cure a patient using a new stem cell therapy—that's huge," says ACT CSO Robert Lanza. "We think we can get very striking results…. The field of stem cells needs a big success."

The company appears to already have overcome many of the hurdles. Preclinical studies in rat and mouse models of retinal degeneration showed that ACT's good manufacturing practice (GMP)-compliant cells brought visual acuity almost back to normal. The cells survived for >8 months without any harmful consequences[5]. The company hopes to be off clinical hold by midsummer and to start dosing patients by fall.

Other research groups are not far behind. An American philanthropist made an anonymous donation to Coffey's group to start the London Project to Cure Blindness, launched in 2007 to drive AMD cell therapy to the clinic within five years. New York–based Pfizer signed on as a partner in 2009, providing an unspecified amount of money, and the company will take the product to market if successful. Coffey hopes to launch a clinical trial next year.

A key decision is which differentiated cell type to use in transplantation. Both ACT and the London Project have settled on RPE cells, which remove shed discs (the photoreceptor outer segments) by phagocytosis and provide photoreceptors with nutrients and oxygen. Undifferentiated stem cells are not an option, because aberrant differentiation might result in disastrous consequences in the eye. Photoreceptors would be attractive, except they still need a healthy RPE to survive, and transplanted cells must reconnect into a diseased neural network. "It's a much more complex replacement strategy," says Coffey. The beauty of RPE, he says, is that it's a single layer of cells in a nonneuronal tissue. Human autologous RPE transplants have already demonstrated that replacement RPE cells can engraft and function. The limitation of RPE transplantation is that it will probably only work in individuals who still have intact photoreceptors.

The other crucial decision is cell source. Inducible pluripotent stem (iPS) cells are one well-publicized option. The 2006 discovery that fibroblasts could be dedifferentiated back into stem cells using only four virally transduced transcription factors bypasses ethical issues. And using the patient's own cells would get around possible immune rejection. The involvement of viral vectors is problematic for human transplantation, and redifferentiation into the cell of choice is difficult, although these technical problems will probably be solved. But "to produce those cells to a clinical grade standard, for an individual, would be immensely expensive," says Coffey. A recent study showed that the iPS cells don't expand as efficiently or uniformly into neuronal cells as hESCs[6]. And reprogramming may only go back so far, with no guarantee that the transplanted cells won't carry the same genetic defects that led to the disease in the first place.

Partially differentiated adult hematopoietic stem cells (HSCs) are another possibility. At least one group has, using a viral vector, genetically reprogrammed mouse HSCs into RPE[7]. But human trials appear to be several years away.

Thus, for now, hESCs remain the closest to human testing. "We've done an awful lot of work on other cell sources of RPE, and to date we can't get anything near as good as what we can produce from human embryonic stem cells," says Coffey. hESCs tend to naturally differentiate into neural ectoderm and, with very little nudging, into functional RPE. And a cell bank can serve as a universal cell source. Lanza says enough cells can be easily and reliably grown to treat millions of people.

Perilous path

But hESCs have their own issues. They're typically cultured initially in media containing mouse feeder cells, raising concerns about contamination with mouse viruses and other pathogens. The final cell product may still contain some undifferentiated ESCs, which could then proliferate abnormally or form teratomas, benign tumors comprised of multiple tissue types. Lanza points out that the mouse cells are taken from animals in closed colonies; that RPE are grown for months without feeders; that everything is done under GMP conditions; and that extensive and redundant testing for pathogens and contaminants takes place throughout the process. And neither ACT nor the London Project have seen any sign of teratomas or other abnormal proliferation in animal models. "We can count 10,000 cells and say there's not a single embryonic stem cell," says Lanza. But only human trials can silence these doubts.

Another theoretical concern is the potential immune response to transplanted allogeneic cells, because the eye is not completely immunoprivileged. In wet AMD, the blood-retinal barrier is breached owing to leaking blood vessels—one reason ACT plans to target the dry form instead. ACT's Stargardt's patients will receive a six-week course of moderate immunosuppression to start. "The hope is that this is just a precaution that is unnecessary," says Lanza.

Delivery method remains an unresolved issue. ACT is injecting its cells subretinally, and the cells then attach to Bruch's membrane and form a new RPE layer. This should work well in Stargardt's, where the Bruch's membrane is intact, and it's one reason ACT has chosen that indication first. (It's also an orphan disease.) But in AMD the Bruch's membrane is damaged. "If you just inject cells in the subretinal space (that is, on the Bruch's membrane) then you don't get good survival or adhesion or polarization of the RPE," says Coffey. Coffey's group plans to implant an engineered patch lined with RPE cells into the back of the eye. That's a more involved procedure than simple injection of the cells, which Lanza points out works in animal models. "They're attaching," he says. But ACT is exploring other approaches just in case.

These pending trials are more important for proving that stem cell therapy works than for any real market impact. A therapy for general use is many years away. But the stem cell trials, along with those testing the many other novel therapies for retinal disease, are an indication of the risks companies are willing to take to succeed in the eye, now that the potential payoff is clear. "The Lucentis results showed us what was possible," says Robert Vitti, Regeneron's head of ophthalmology. "Sometimes you need that science to think beyond the present."

Ken Garber, Ann Arbor, Michigan

1. Brown, D.M. *et al. N. Engl. J. Med.* **355**, 1432–1444 (2006).
2. Jo, N. *et al. Am. J. Pathol.* **168**, 2036–2053 (2006).
3. Maguire, A.M. *et al. Lancet* **374**, 1597–1605 (2009).
4. Nishijima, K. *et al. Am. J. Pathol.* **171**, 53–67 (2007).
5. Lu, B. *et al. Stem Cells* **27**, 2126–2135 (2009).
6. Hu, B.Y. *et al. Proc. Natl. Acad. Sci. USA* **107**, 4335–4340 (2010).
7. Sengupta, N. *et al. Mol. Ther.* **17**, 1594–1604 (2009).

Corrected after print 15 April 2010.

Erratum: Biotech in a blink

Ken Garber

Nat. Biotechnol. **28,** 311–314 (2010); **published online 8 April 2010; corrected after print 15 April 2010**

In the version of the article originally published, Michael Tolentino was misquoted to the effect that bevasiranib had been shown to persist indefinitely in post-mitotic cells. Tolentino actually stated that the RNA-induced signaling complex persists. The error has been corrected in the HMTL and PDF versions of the article.

Selling out

Joel F Martin

Preferred stock preferences bite into an entrepreneur's returns. What can you do to put venture capitalists on a diet?

It's finally payday! You've spent years building your company, clawed your way through financing after financing, and perhaps suffered the slings and arrows of risky, expensive clinical trials. Now a large pharmaceutical company has offered to buy you out. It has taken more money and time than expected, but you and your team still own a respectable percentage of the company, so it's time to open the champagne!

This is when your trusted corporate counsel pulls you aside for a little talk. "You see," she or he says, "there is ownership and there's *ownership*, and they aren't equal." Your attorney is talking about the types of stock owned by employees (common stock) and investors (preferred stock) and the differences between the two.

On the basis of my own experiences as a venture capitalist, investors always understand these differences; entrepreneurs almost never do. In this article, I explain the differences and how to prepare your company (and yourself) for the best possible outcome.

Forewarned is forearmed

Especially in today's difficult fund-raising environment, companies may be forced to accept stringent terms to attract capital, and that means more money for investors and less money for you at exit. Back when the initial public offering (IPO) market was hot, the differences between classes of stock were modest and could often be ignored. That's because investors are forced to convert to common stock as part of an IPO and all of the preferential treatment accorded to preferred stock disappears. But today, IPOs are rare and acquisition is the preferred route of exit. This creates complex layers of expensive stock preferences, and this is bad news for founders.

There are two key financial preferences that differentiate the payouts between common and preferred shareholders (**Box 1**). The first is the liquidation preference, and the second is the participation right.

There are good reasons why preferences exist. Consider the simple case in which investors pay $10 million for a 50% ownership stake in a company. Suppose that the company is then sold for $10 million. On a pure ownership basis, the common shareholders (namely, management and founders) and the investors would each receive $5 million, resulting in a $5 million gain for common shareholders and a $5 million loss for investors. Here, wealth is simply transferred to management. Investors guard against this through liquidation preferences that guarantee that they get their money back before anyone else gets a penny.

Today, liquidation preference multiples greater than one are increasingly common, particularly in later-stage financings. This means that investors get back a multiple on their investment before money goes to common shareholders. In high-tech financings, such multiples are being imposed as early as first rounds, and biotech terms may not be far behind. Sometimes, investors receive two or three times their investment, or even higher multiples, before common shareholders receive any proceeds whatsoever.

There's a similar trend in participation rights, with 64% of all financings having full participation, providing investors with unlimited participation in the gains (**Box 1**).

High liquidation multiples are common in venture capital 'bridge loans', which are usually put in place to bridge a company to an equity financing, a merger or an acquisition transaction. If sold while the loan is in place—not an unusual happenstance—the company's liquidation preferences come into play and the investors enjoy an up-front multiple on their investment. Call options (referred to as warrants) further increase the returns to the loan investors (**Box 2**).

There are three things you as an entrepreneur can do to mitigate the effect of preferences. First, understand the capital structure of your company and any pending financings. Second, communicate openly with your investors and your board about any inequities, unrealistic terms and disincentives to management. Third, limit

Box 1 The rights of investors

Almost all venture capital financings contain at least one liquidation preference. The main types of these deal terms are liquidation rights and participation rights. The former gives the preferred shareholders the right to get their money back first in the event of the sale, merger or winding down of a company. This preference can have a multiplier that ranges from zero to several times the invested amount. In contrast, a participation right enables preferred shareholders (investors) to continue to collect their respective 'ownership' of the remaining proceeds alongside the common shareholders, even after they have claimed their liquidation preferences. The participation right can be capped at a specific multiple or can be unlimited.

A comparatively innocuous set of terms is a liquidation preference that allows investors to get their money back first with no subsequent participation rights. Under those terms, investors simply take the larger of two options—either getting their money back or converting to common stock and taking their respective ownership percentage.

A more punishing extreme, formerly imposed on companies in trouble and now more broadly, occurs when investors get a multiple of their capital back first and then get their percentage on top of that.

Joel F. Martin is president and CEO of Altair Therapeutics, San Diego, California, USA.
e-mail: jfmartin@altairthera.com

Box 2 Dividends and warrants

Some financings provide for cumulative preferred stock dividends of 8%–9% for the preferred shareholders but not for the common shareholders. These dividends effectively increase the liquidation preference multiple over time. With many companies attaining liquidity after seven or more years, such dividends can amount to a liquidation premium of 1.5 times the investment, or more,for some series of preferred stock. Both the layers of preferences and the dividends can make for an unwelcome surprise for a CEO who thinks he faces a liquidation preference that is one times the investment.

Another preferred stock sweetener is called a warrant. It is a type of stock option that allows investors to buy additional stock at a preset price in the future. Warrants are similar to employee stock options but typically have more complex terms. They are frequently imposed in difficult financing situations as a means to boost potential returns to investors. Warrants further dilute common shareholders during a successful exit.

the amount of capital you raise to the minimum needed to get the work done.

Capital structure

It takes skill and hard work to thoroughly understand capital structures. Entrepreneurs are usually more interested in products and science than financing documents, and laziness may prevail over diligence. Older companies with multiple rounds of financing are particularly hazardous, with a punitive overhang of preferences and payout formulas that run into many pages of complex legalese.

The accumulated preferences can be especially punitive when companies have been through difficult down rounds, known as 'cram down' or 'washout' rounds in the colorful vernacular of venture capitalists. In these rounds, a company's stock price is significantly reduced and the ownership of earlier investors is diluted proportionately. The problem is that venture capitalists leading a washout financing may allow earlier investors (with whom they will be doing future deals) to retain their liquidation preferences to preserve goodwill and discourage retaliation. Alas, the overhanging liquidation preferences can completely wipe out entrepreneurs.

The solution is to understand your capital structure and its ramifications. Start by working with your corporate counsel and an experienced financing expert to build a spreadsheet model. Don't count on your counsel to flag potential problems without prompting. They work with your investors on many deals (compared to your one deal), and they are not looking to bite the hand that feeds them. However, they will walk you through the process if asked—so by all means, ask.

Talk, talk, talk

If there are inequities in your capital structure, you need to martial your resources and rectify the issues by communicating with your board and investors. Key points to address are comparability to industry norms, adverse incentives to management and fairness. Here a compensation consultant can work wonders in validating appropriate comparables and norms. Remember, your investors want you incentivized to maximize their returns. If your incentives point in the wrong direction, investors should be interested in aligning your interests with theirs. The penalties to investors for not aligning interests can be severe. For example, management may forego an attractive opportunity to be acquired (because of overhanging preferences) in favor of a highly speculative effort to go public (because it wipes out those preferences).

There are several ways to encourage alignment, including appealing to investors' common sense—if there's nothing in it for you, there's likely nothing in it for them. Additionally, incoming investors can be very powerful allies. They have fresh enthusiasm, their preferences are senior to those of the other investors and they have a strong desire to see you motivated. Legacy investors are prone to protecting preferences on investments that may really be sunk costs. Understanding the circumstances and motivations of each set of investors is the key to any capital structure renegotiation. Antagonistic incentives between different classes of investors can produce amazing fireworks at board meetings but will be detrimental to your company if left unresolved.

A variety of mechanisms can be used to align the incentives of investors and management. The most routine is to issue more common stock, though this doesn't solve the problem of overhanging preferences. The second is a carve-out provision that sets aside a pot of money for the management. The amount of the carve out is highly variable (for example, 2%–10% or more) depending on the negotiating skill and leverage of management. Carve outs are typically negotiated in the heat of an acquisition, often when management discovers they are getting skunked. Carve outs usually favor the company's most senior management because of their importance in closing the deal. Other employees, and non-employees such as scientific founders, may be left out. These shareholders then find their payouts diminished not only by the preferred stock preferences but also by the carve out. There are other mechanisms for aligning interests, ranging from a set aside for common shareholders to converting old preferred stock to common stock as part of a new financing.

The worst time to renegotiate a capital structure is when there is an acquisition offer on the table. Management has already done its job and investors are focused on keeping just a few key people that the acquirer wants to retain. Any leverage management had is gone, leading to two deleterious consequences. First, the pot of money set aside for the common stock is minimized. Second, a key person may realize his importance to the deal and subsequently extort a disproportionate piece of the pie. I have seen this happen and can vouch that it causes infighting and tremendous unhappiness, and it can even kill a deal. Generally, the sooner that structural issues are addressed, the better. There are particularly propitious times from a leverage perspective. These include successful times (the negotiation is easier when investors are happy), new financings (when the incoming investors are most aligned with management) and management changeover (when new management has unique leverage in cleaning up structural problems).

Pass on the peas

The simplest way to mitigate preferences is to obtain a high multiple on invested capital. To get that multiple, you can either increase the numerator (the price) or decrease the denomi-

Table 1 Return on total invested capital[a]

Total invested (millions)	0–20	20–40	40–60	60–80	80–100	>100
Average investment (millions)	11.4	29.7	50.5	72.0	87.7	134
Average return value (millions)	73.0	132.4	160.3	196.2	166.0	202.3
Average return multiple	6.4	4.5	3.2	2.7	1.9	1.5
Return multiple standard deviation	2.5	3.2	2.4	1.3	0.8	0.6
Number of exits	9	15	21	15	10	19

[a]Statistics derived from Booth, B.L. *Nat. Biotechnol.* **25**, 853–857, 2007 and pooled by investment amount rather than return multiple.

nator (the amount invested). The acquisition price (numerator) is largely beyond your control, but the money you raise (denominator) is not. The trick is capital efficiency. Remember, every dollar taken in financing is a dollar added to overhanging liquidation preferences. Larding up on unnecessary cash inflates not only liquidation preferences but also the 'post-money valuation' of the company, or the company's value after investment. This makes it hard to find an acceptable exit for either the investors or the entrepreneurs.

The notion of taking less money than offered is antithetical to many in biotech. Some luminary first-generation biotech venture capitalists will say that "the time to take the peas is when they are passing the peas." Namely, take the money when it's on the table. That was probably sage advice in the days when biotech was focused on inexpensive research (compared to clinical development) and IPOs were the exit of choice. Today, venture financings are much larger. According to PricewaterhouseCoopers' (New York) MoneyTree Report, the average biotech financing ballooned from $5.7 million in 1998 to $10.8 million in 2007 (based on the cash infusion per tranche rather than the total round size). Over that period, larger venture fund sizes incentivized investors to put more money to work per financing.

Usually, companies raise prudent sums to support clinical trials or secondary programs that have positive net present value even at the high cost of capital imposed by venture financings. Unfortunately, I've seen many instances in which companies take down too much capital simply because it is available. A company may wish to build a cash cushion, fund excessively speculative projects, build manufacturing infrastructure or simply increase headcount when outsourcing would suffice. Companies may raise capital for early-stage backup programs to mitigate the risk of failure for later-stage lead programs, even when such programs wouldn't justify investment on their own. Sadly, in the event that the first program fails, management—at least the current management anyway—may not be around to advance a second program.

Taking too much cash does a disservice on two counts. The first is the penalty exacted by liquidation preferences as we have discussed. The second is that, historically, more capital correlates with lower multiples on invested cash. Coupled, these two factors can erase much of the economic incentive that has traditionally spurred entrepreneurship.

An inverse correlation between return multiples and increasing capitalization is shown in **Table 1**. There is a clear take-home message from these data. The return multiples declined from 6.4 times to 1.5 times the investment as the capital raised grew from $20 million to $100 million. Lower multiples exacerbate all of the problems described above (**Fig. 1**). (The model explicitly assumes an acquisition in which the preferences come into play.)

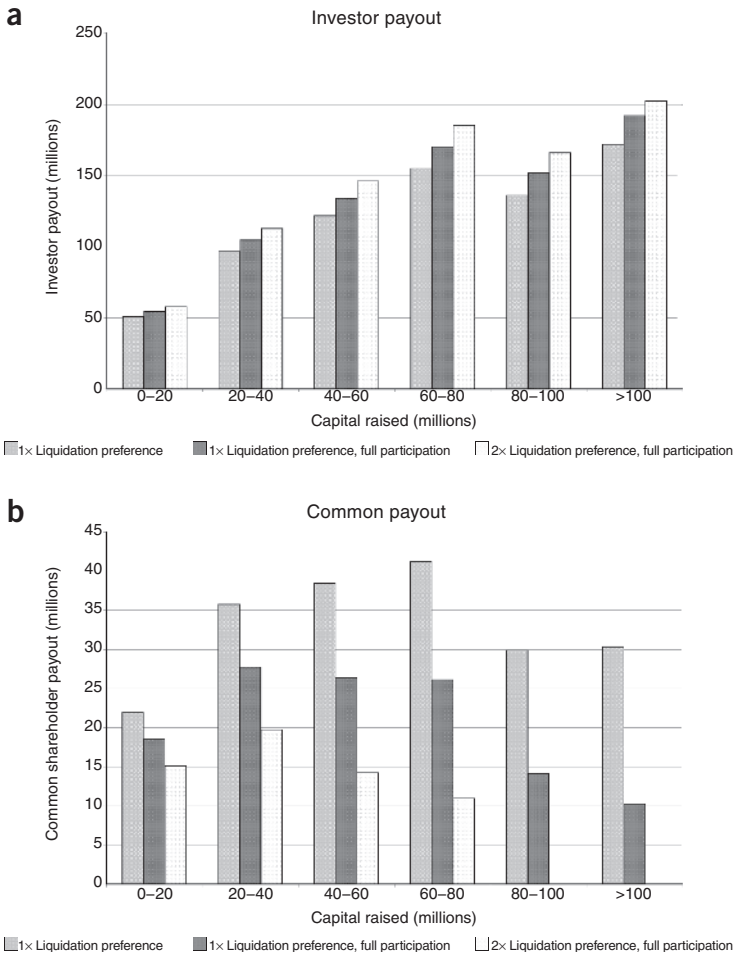


Figure 1 Returns. (**a**) Investor payout against capital raised. (**b**) Common shareholder payout against capital raised

Regardless of the specific capital structure, the money returned to common shareholders peaks at less than $80 million of invested capital. The more stringent the preference terms, the earlier the peak, even when we ignore all the other things that cut into common shareholder returns (for example, payments to bankers, debt repayment, warrants and cumulative preferred dividends).

The model relies on several assumptions, particularly the tabulated return data and an idealized capital structure. However, it probably understates the problems of raising too much capital. Acquisition prices have dropped sharply in the past year (particularly up-front cash payments), so the model is probably optimistic. The lesson is twofold: first, raise capital prudently with a sober perspective on probable outcomes, and second, if you are coming into a company that has already raised a lot of cash, understand the consequences of the exiting capital structure; it may yet be possible to change the deal.

A balanced meal

Entrepreneurs need to keep the totality of their businesses in mind. Tough financing terms and large capital requirements interact to punish entrepreneurs for the life of their companies. It's critical to understand what you're getting into and what it means down the road. Oddly enough, it may be better to take a lower valuation and favorable terms than a high valuation and punitive terms. It pays to do the math! Entrepreneurs need to understand the implications of their capital structures and rectify inequities before they become an issue. Furthermore, they need to analyze projects such that they provide positive returns with a cost-of-capital that includes the full burden of preferences. To stay aligned, all shareholder classes should strive for fair, transparent terms, an understanding of the implications of those terms and an eye toward capital efficiency.

Making the most of GM potatoes

To the Editor:

The recent approval of the Amflora potato by the European Union (EU)—the EU's first registration of a genetically modified (GM) potato in 12 years—has garnered considerable media attention and public controversy. Amflora (EH92-527-1) is a GM potato produced by BASF (Ludwigshafen, Germany) that lacks amylose and instead contains amylopectin (>98%) as the predominant starch[1,2]. Amylose ordinarily has to be removed to allow the industrial use of potato starch. Thus, Amflora is a highly suitable source for technical applications, such as paper, adhesive and textile production. Supporters of the technology welcome the approval, which has taken 13 years, and consider it a regulatory milestone, at least for GM potatoes. Opponents are afraid that it heralds the opening of the regulatory floodgates for more transgenic varieties. Accepting the view that the use of GM technology should be based on careful case-by-case consideration[3], I see two key issues that may not be trivial and should be discussed, as they generate most of the public concern.

Amflora contains a gene encoding neomycin phosphotransferase II (NptII) that confers kanamycin resistance, and critics argue that this antibiotic resistance gene could escape via the food chain or horizontal transfer into ecosystems. The antibiotic marker is there as a selection gene and was needed to introduce the antisense construct that blocks amylose production by targeting granule-bound starch synthase (GBSS). Although horizontal transfer of transgenic traits into ecosystems is not well accepted by the scientific community, the possibility that it could happen cannot completely be excluded[4,5]. One way around this possibility would be simply to remove the selection gene *NptII* (e.g., using the Cre/loxP system[6]) in future generations. This would represent an additional burden for the breeders of GM potatoes, but it would also facilitate a second round of transformation, if needed. Most importantly of all (scientific concerns aside), an Amflora derivative lacking kanamycin resistance would have much improved public acceptance. In fact, if in subsequent years

You say potato, I say Amflora. BASF recently received approval from the EU to market its GM potato engineered with reduced amylase content via an antisense construct targeting granule-bound starch synthase.

Amflora were to inadvertently end up in the human food chain through admixture with potatoes grown for human consumption, the potential health risk would be diminished, as removal of the antibiotic marker would leave only the transgene, an antisense construct, which lowers the expression of an endogenous GBSS gene but has no protein-coding potential on its own.

A second concern of opponents of Amflora potatoes relates to the possibility of transgene dissemination to other potatoes. Such spreading of the transgene is unlikely, as potato transgene movement by pollen is very limited[7] and escaped wild-type potatoes have rarely been observed in Europe[8]. Nevertheless, these arguments again cannot be dismissed completely, as in our global world, dissemination in the long term cannot be excluded and may even be likely. As commercial potato production, especially that using GM varieties, does not require sexual reproduction, it seems reasonable to carry out gene manipulation in potato varieties in which the genes for sexual reproduction have been permanently deleted. Such a strategy would probably negate concerns relating to the spread of potato transgenes into the ecosystem.

In conclusion, with the addition of these two safety features, GM potatoes could become the standard for other transgenic crops, particularly in the European market, where outcrossing or admixture of GM crops with conventional varieties remains such a hot-button issue. Such crops would represent low-risk GM varieties, which possibly could be cleared through approval authorities in a more timely manner[9]. For example, GM potatoes resistant to potato late blight[10] that are currently being generated (http://www.gmo-safety.eu/en/potato/plant_diseases/462.docu.html) would probably be more palatable to both the public and regulatory authorities if selection markers were removed and sexual reproduction were irreversibly blocked. Who knows: as late-blight disease caused by the pathogen *Phytophthora infestans* is a serious problem in the farming of organic potatoes[11], sterile GM potatoes resistant to the phytopathogen might become accepted even in the organic farming community[12].

COMPETING FINANCIAL INTERESTS
The author declares no competing financial interests.

Gerhart U Ryffel

Institut für Zellbiologie (Tumorforschung), Universitätsklinikum Essen, Universität Duisburg-Essen, Germany.
e-mail: gerhart.ryffel@uni-due.de

1. EFSA. *The EFSA Journal* **323**, 1–20 (2006).
2. EFSA. *The EFSA Journal* **324**, 1–20 (2006).
3. Arntzen, C.J., Coghlan, A., Johnson, B., Peacock, J. & Rodemeyer, M. *Nat. Rev. Genet.* **4**, 839–843 (2003).
4. Lemaux, P.G. *Annu. Rev. Plant Biol.* **59**, 771–812 (2008).
5. Lemaux, P.G. *Annu. Rev. Plant Biol.* **60**, 511–559 (2009).
6. Cuellar, W. *et al. Plant Mol. Biol.* **62**, 71–82 (2006).
7. McPartlan, H.C. & Dale, P.J. *Transgenic Res.* **3**, 216–225 (1994).
8. Lauber, K. & Wagner, G. *Flora des Kantons Bern: 1836 Farbfotos der wildwachsenden Blüten- und Farnpflanzen; Artbeschreibung und Bestimmungsschlüssel* (Paul Haupt, Berne, Switzerland, 1991).
9. Rommens, C.M. *J. Agric. Food Chem.* **55**, 4281–4288 (2007).
10. Song, J. *et al. Proc. Natl. Acad. Sci. USA* **100**, 9128–9133 (2003).
11. Dupuis, B., Rolot, J.L., Stilmant, D., Labbe, V. & Laguesse, L. *Commun. Agric. Appl. Biol. Sci.* **72**, 353–359 (2007).
12. Ronald, P.C. & Adamchak, R.W. *Tomorrow's Table: Organic Farming, Genetics and the Future of Food* (Oxford Univ. Press, New York, 2008).

Peer-reviewed surveys indicate positive impact of commercialized GM crops

To the Editor:

The benefits of genetically modified (GM) crops continue to be disputed, despite rapid and widespread adoption since their commercial introduction in the United States and Canada in 1995. Last year, 14 million farmers in 25 countries grew GM crops commercially, over 90% of them small farmers in developing countries[1]. Farmer surveys are a valuable measure of the impact of GM crops. These surveys estimate the technology's performance as it is incorporated into farmer practices, given constraints on time, access to information, differing levels of risk aversion and other factors. This analysis summarizes results from 49 peer-reviewed publications reporting on farmer surveys that compare yields and other indicators of economic performance for adopters and non-adopters of currently commercialized GM crops. The surveys cover GM insect-resistant and herbicide-tolerant crops, which account for >99% of global GM crop area[1]. Results from 12 countries indicate, with few exceptions, that GM crops have benefitted farmers. The benefits, especially in terms of increased yields, are greatest for the mostly small farmers in developing countries, who have benefitted from the spillover of technologies originally targeted at farmers in industrialized countries.

Of 168 results comparing yields of GM and conventional crops, 124 show positive results for adopters compared to non-adopters, 32 indicate no difference and 13 are negative. By far the largest numbers of results comparing yields of adopters and non-adopters come from India and the United States, which account for 26% and 23% of the results, respectively (**Table 1**). An annotated bibliography of results for yield, costs and economic performance, and a description of the methodology used in this analysis, can be found in **Supplementary Tables 1** and **3**.

The results for yields indicate that farmers in developing countries are achieving greater yield increases than farmers in developed countries (**Table 2**). The average yield increases for developing countries range from 16% for insect-resistant corn to 30% for insect-resistant cotton, with an 85% yield increase observed in a single study on herbicide-tolerant corn. On average, developed-country farmers report yield increases that range from no change for herbicide-tolerant cotton to a 7% increase for herbicide-tolerant soybean and insect-resistant cotton. The first wave of GM crops to be commercialized has embodied traits intended to improve pest management and therefore reduce or eliminate losses from insect damage or weed competition. These technologies do not raise yield potential, but they can improve yields substantially owing to improved pest management. Where conventional weed- and insect-control technologies were lacking because of inherent limits to the effectiveness of available conventional pest-management options or limited access to conventional control methods, yields would be expected to increase. These conditions may be more common in developing countries.

As the most frequently studied case, GM insect-resistant cotton (*Bacillus thuringiensis* (*Bt*) cotton) in India provides examples of both the highest yield increases observed as well as several of the negative results. The largest yield increases found in this review are reported for *Bt* cotton in India, where surveys show yield increases of up to 150%. Of the negative results, six are for the first year of commercialization of *Bt* cotton in India, and the rest of the negative results are from developed countries in the first few years of commercialization.

The results show the variability of benefits from region to region and year to year. A survey of Indian cotton farmers in crop harvest years 2005–2006 through 2007–2008 showed that *Bt* cotton growers in Gujarat had larger yield improvements than their counterparts in Maharashtra, with the former obtaining 82–150% greater yields, whereas the latter obtained only 24–40% higher yields. Smallholders in KwaZulu Natal, South Africa, who were surveyed from harvest years 1998–1999 to 2000–2001 reported a yield benefit associated with *Bt* cotton between 56% and 85%, which is attributed to variable weather conditions and pest pressure from year to year. It is important to note that the analysis of yield differences is complicated by differences in yield potential and other characteristics of background germplasm that may differ between the varieties that are available with and without the engineered trait. For example, the first *Bt* cotton varieties to be approved for commercialization in India had been in the regulatory pipeline for several years, during which time conventional breeding had continued to produce varieties with superior yields and disease resistance. These earliest official varieties were known to be susceptible to wilt when subjected to early moisture stress, which may have driven the negative results observed

Table 1 Number and direction of results comparing yields of GM adopters to those of non-adopters, by country

Country	Positive	Neutral	Negative	Total
Developed countries	36	18	7	61
Australia	0	2	2	4
Canada	7	0	1	8
Spain	3	6	0	9
United States	26	10	4	40
Developing countries	88	13	6	107
Argentina	5	1	0	6
China	15	0	0	15
Colombia	4	1	0	5
India	35	2	6	43
Mexico	2	0	0	2
Philippines	5	2	0	7
Romania	2	0	0	2
South Africa	20	7	0	27
Total	124	32	13	168

Positive and negative directions refer to a comparison of GM to conventional crops.

in Andhra Pradesh in the first year of commercialization[2,3]. In the early years of commercialization, it is likely that the technology would not be available in the highest-yielding background varieties or in varieties that are most suited to the growing conditions in all areas.

Profitability is an important measure that complements data on yields, as even a technology that does not necessarily increase yields can improve a farmer's bottom line if it reduces costs. In addition to yields, many of the surveys reviewed here also look at changes in costs and various measures of farm economic performance. In all but one case reviewed, the cost of seeds (including any technology fees) rose. However, this was offset by decreases in pesticide costs, which were found in all but 12 cases.

Looking across all measures of economic performance, the results are also overwhelmingly positive. Gross margins are most commonly reported, but the variable costs that are included in these calculations vary greatly from study to study. Of the 98 results in our survey of the peer-reviewed literature that compare the economic performance of GM crops to their conventional counterparts, 71 indicate a positive impact, 11 neutral and 16 negative (**Fig. 1**).

For GM herbicide-tolerant crops, 12 of 17 results show a positive impact on economic performance, whereas 4 results show no difference and 1 result shows a negative impact. One might expect more results showing positive impacts of GM herbicide-tolerant crops on economic performance, particularly as GM herbicide-tolerant crops have been more widely adopted (on 62% of global GM crop acreage in 2009) than GM insect-resistant crops around the world[1]. This may be due to cost savings associated with GM herbicide-tolerant crops that are not included in a traditional accounting of costs. In a study of glyphosate-tolerant soybean in the United States, nonmarket valuation techniques were used to estimate 'non-pecuniary' convenience benefits, such as management-time savings and flexibility, at $12 ha^{-1} (ref. 4).

For GM insect-resistant crops, 59 of 80 results indicate improved economic performance, 7 results are neutral and 14 results are negative. On the positive side, some of the most striking results come from *Bt* cotton growers in South Africa and China. Negative results are in *Bt* corn in the United States and *Bt* cotton in Australia, China, Colombia, India and South Africa (though positive results are reported in each of these cases as well). Some of the negative results may be explained by year-to-year variation in pest pressure and technology pricing. Furthermore, similar to GM herbicide-tolerant crops, farmers may value intangible benefits of GM insect-tolerant crops. A survey of US corn farmers found that non-pecuniary benefits (handling and labor-time savings, human and environmental safety, reduced yield risk, equipment cost savings and better standability) of GM insect-resistant corn were valued at $10.32 ha^{-1} (ref. 5).

In addition to economic indicators of performance, many surveys also look at indicators of the environmental impact of GM crops, specifically changes in tillage practices for GM herbicide-tolerant crops and changes in pesticide use for GM insect-resistant and herbicide-tolerant crops (**Supplementary Tables 2** and **3**). For GM herbicide-tolerant crops, two surveys, for soybeans in Argentina and the United States, report decreases of 25–58% in the number of tillage operations[6,7]. There are no results indicating an increase in tillage for adopters of GM herbicide-tolerant crops. These results reinforce observations of wider adoption of conservation tillage practices since the introduction of GM herbicide-tolerant crops[8–10].

For insect-resistant crops, 45 results show decreases in the amount of insecticide or number of insecticide applications, or both, used on *Bt* crops compared with conventional crops in Argentina, Australia, China, India and the United States. The reductions range from 14% to 75% in terms of amount of active ingredient and from 14% to 76% for the number of insecticide applications. A small sample survey in South Africa found a reduction in the number of insecticide sprays in one of two years studied and an insignificant difference in the other year. There are no results indicating an increase in insecticide use for adopters of GM insect-resistant crops.

The above measures of changes in insecticide use are imperfect, in that they do not indicate the relative toxicity of insecticides used and therefore the human health and environmental impacts associated with the different insecticides that might be used on *Bt* and conventional crops. Researchers have used various approaches to give further insight into the implications of these reductions in insecticide use. In Argentina, a survey showed that the amount of insecticide in all toxicity classes was reduced in *Bt* compared with conventional cotton plots[11]. Three years of survey data for *Bt* cotton in South Africa were combined with ratings of relative toxicity and persistence to calculate a biocide index, which showed substantially lower values for insecticides used on *Bt* compared with conventional cotton[12]. Survey results in China indicate a reduction in the percentage of farmers reporting headaches, nausea, skin pain or digestive problems after applying pesticides associated with adoption of *Bt* cotton[13], though some of the health benefits observed in the early years of adoption may

Table 2 Average impact on yield, by technology, for developed and developing countries

Technology	Difference in yield (%)	Number of results	Minimum (%)	Maximum (%)	Standard error of the mean (%)
Developed countries	6	59	–12	26	1.0
Herbicide-tolerant cotton	0	6	–12	17	3.8
Herbicide-tolerant soybean	7	14	0	20	1.7
Herbicide-tolerant and insect-resistant cotton	3	2	–3	9	5.8
Insect-resistant corn	4	13	–3	13	1.6
Insect-resistant cotton	7	24	–8	26	1.9
Developing countries	29	107	–25	150	2.9
Herbicide-tolerant corn	85	1			
Herbicide-tolerant soybean	21	3	0	35	11
Insect-resistant corn	16	12	0	38	4
Insect-resistant corn (white)	22	9	0	62	6.9
Insect-resistant cotton	30	82	–25	150	3.5

Yield difference for adopters was calculated as (GM yield – conventional yield)/conventional yield, averaging yields across surveys, geographies, years and methodologies. The difference in the number of results reported in **Tables 1** and **2** is due to two results reported as 'positive' with no numerical value. A two-tailed *t*-test shows a significant difference between the average yields of developed and developing countries ($t = 7.48$, df = 134, $P < 0.0005$).

have been eroded by increased spraying for secondary pests[14]. Researchers in South Africa have shown an inverse relationship between the number of local hospital admissions classified as related to cotton growing and the adoption of *Bt* cotton[15].

Few surveys have captured changes in herbicide use with GM herbicide-tolerant crops, perhaps because the impact of GM herbicide-tolerant crops has largely been a switch between herbicides that are applied at different rates, and therefore change in the amount of herbicide used is a poor indicator of environmental impact. Three surveys report changes in herbicide use, showing changes that range from a decrease of 38% to an increase of 108% in the total amount of herbicide used, and an insignificant change in the number of herbicide applications[6,16,17]. The environmental impact of these shifts is better understood by looking at the environmental characteristics of the herbicides. Two of the studies above extend their analysis by applying environmental indicators to observed changes in herbicide use. The aggregate pesticide leaching potential for GM herbicide-tolerant cotton in North Carolina was 25% lower than that of conventional cotton[16]. Reductions of 83% and 100% in the use of herbicides in toxicity classes II and III, respectively, were found in GM herbicide-tolerant soybeans in Argentina, with a corresponding increase of 248% in the use of less toxic class IV herbicides[6]. Some of the environmental benefits that have come with the use of more environmentally benign herbicides may be eroded with the development of glyphosate-resistant weeds, although few data now exist upon which to draw any conclusions.

Several surveys address the question of whether GM crops are benefitting small farmers in developing countries through direct comparisons of outcomes for farmers with different-sized land holdings or by documenting the impacts on small farms. Four surveys from China, Colombia and South Africa make direct comparisons of yields, gross margins or both for farmers with different-sized operations. The surveys indicate that the smallest farmers benefitted most in South Africa and China[15,18,19]. Results from Colombia were mixed[20]. Five studies have shown improvements in economic performance for farmers with <10 ha in China, Colombia, Mexico, India and South Africa[13,20–23]. One explanation of the favorable outcomes for smallholders is the risk-reducing nature of the technology,

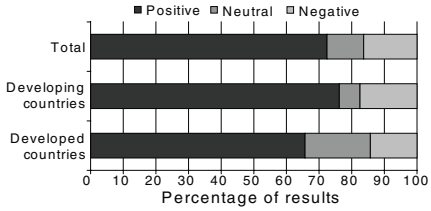


Figure 1 Results by direction of change in economic performance (GM – conventional). A χ^2 test shows a significant difference in the proportion of positive results for developed and developing countries ($\chi^2 = 0.68$, df = 1, $P = 0.41$).

whereas non-adopters with comparably sized land holdings are particularly vulnerable as yields fluctuate from year to year[15].

The accumulated evidence from farmer surveys on the performance of GM crops helps to explain the widespread popularity of the technology in several regions of the world. The surveys reviewed here reflect a wide variety of conditions in terms of environment, pest pressure, farmer practices, social context, intellectual property rights and institutional arrangements. Given this diversity of conditions, it is striking that the results are so consistently positive. Even so, these results cover less than half of the countries currently growing GM crops and are sparse for some already widely adopted technologies, such as GM herbicide-tolerant corn and canola. Furthermore, GM crops have been grown for only 14 years—fewer for those countries that were not among the first adopters—a relatively short period for assessing the long-term impact of any technology. In some cases, results reflect a single growing season, which may not be an adequate basis for judging the sustainability of the technology's impact. Nevertheless, the window of opportunity for directly comparing the outcomes of adopters and non-adopters has closed where adoption rates are very high, and different methods of impact assessment will now be required.

Of interest in the future will be the assessment of the impacts of stacked traits, incorporating a combination of traits, which already represent over 28% of total global GM crop acreage[1] but have been studied by only two surveys. Also of interest will be the assessment of farmers' experiences with GM crop technologies created specifically to address the most pressing constraints of developing-country farmers, such as technologies being developed in cassava, cowpea and rice, as those reach the commercialization stage.

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

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

Janet E Carpenter

PO Box 1008, Boylston, Massachusetts, USA.
e-mail: janet.e.carpenter@gmail.com

1. James, C. *Global Status of Commercialized Biotech/GM Crops: 2009—The First Fourteen Years, 1996 to 2009* (International Service for the Acquisition of Agri-Biotech Applications, Ithaca, New York, USA, 2010).
2. Sahai, S. & Rahman, S. *Econ. Polit. Wkly.* 26 July 2003, 3139–3141.
3. Herring, R.J. *Int. J. Mul. Res. Approaches* **2**, 145–159 (2008).
4. Marra, M. & Piggott, N. in *Regulating Agricultural Biotechnology: Economics and Policy* (eds. Just, R.E., Alston, J.M. & Zilberman, D.) 145–178 (Springer Science and Business Media, New York, 2006).
5. Alston, J.M., Hyde, J., Marra, M. & Mitchell, P.D. *AgBioForum* **5**, 71–84 (2002).
6. Qaim, M. & Traxler, G. *Agric. Econ.* **32**, 73–86 (2005).
7. Marra, M.C., Piggott, N.E. & Carlson, G.A. *The Net Benefits, Including Convenience, of Roundup Ready Soybean: Results From a National Survey.* Technical Bulletin 2004–3 (NSF Center for IPM, Raleigh, North Carolina, USA, 2004).
8. Fawcett, R. & Towery, D. *Conservation Tillage and Plant Biotechnology: How New Technologies Can Improve the Environment By Reducing the Need to Plow* (Conservation Tillage Information Center, West Lafayette, Indiana, USA, 2003).
9. Givens, W.A. *et al. Weed Technol.* **23**, 150–155 (2009).
10. Trigo, E.J. & Cap, E.J. *Ten Years of Genetically Modified Crops in Argentine Agriculture* (ArgenBio, Buenos Aires, 2006).
11. Qaim, M. & de Janvry, A. *Environ. Dev. Econ.* **10**, 179–200 (2005).
12. Bennett, R., Ismael, Y., Morse, S. & Shankar, B. *J. Agric. Sci.* **142**, 665–674 (2004).
13. Huang, J., Hu, R., Fan, C., Pray, C.E. & Rozelle, S. *AgBioForum* **5**, 153–166 (2002).
14. Wang, Z. *et al. Agric. Sci. China* **8**, 101–105 (2009).
15. Bennett, R. *et al. J. Dev. Stud.* **42**, 662–677 (2006).
16. Wossink, A. & Denaux, Z.S. *Agric. Syst.* **90**, 312–328 (2006).
17. Cattaneo, M.G. *et al. Proc. Natl. Acad. Sci. USA* **103**, 7571–7576 (2006).
18. Ismael, Y., Bennett, R. & Morse, S. *AgBioForum* **5**, 1–5 (2002).
19. Pray, C., Ma, D., Huang, J. & Qiao, D. *World Dev.* **29**, 813–825 (2001).
20. Zambrano, P., Fonseca, L.A., Cardona, I. & Magalhaes, E. in *Biotechnology and Agricultural Development: Transgenic Cotton, Rural Institutions and Resource-Poor Farmers* (ed. Tripp, R.) 168–199 (Routledge, London, 2009).
21. Qaim, M., Subramanian, A., Naik, G. & Zilberman, D. *Rev. Agr. Econ.* **28**, 48–58 (2006).
22. Morse, S., Bennett, R. & Ismael, Y. *AgBioForum* **10**, 44–50 (2007).
23. Traxler, G., Godoy-Avila, S., Falck-Zepeda, J. & De Jesus Espinoza-Arellano, J. in *The Economic and Environmental Impacts of Agbiotech: A Global Perspective* (ed. Kalaitzandonakes, N.) 183–202 (Kluwer Academic/Plenum Publishers, New York, 2003).

A global map of human gene expression

To the Editor:

Although there is only one human genome sequence, different genes are expressed in many different cell types and tissues, as well as in different developmental stages or diseases. The structure of this 'expression space' is still largely unknown, as most transcriptomics experiments focus on sampling small regions. We have constructed a global gene expression map by integrating microarray data from 5,372 human samples representing 369 different cell and tissue types, disease states and cell lines. These have been compiled in an online resource (http://www.ebi.ac.uk/gxa/array/U133A) that allows the user to search for a gene of interest and find the conditions in which it is over- or underexpressed, or, conversely, to find which genes are over- or underexpressed in a particular condition. An analysis of the structure of the expression space reveals that it can be described by a small number of distinct expression profile classes and that the first three principal components of this space have biological interpretations. The hematopoietic system, solid tissues and incompletely differentiated cell types are arranged on the first principal axis; cell lines, neoplastic samples and non-neoplastic primary tissue–derived samples are on the second principal axis; and nervous system is separated from the rest of the samples on the third axis. We also show below that most cell lines cluster together rather than with their tissues of origin.

The widely used *GNF Gene Expression Atlas*[1,2] includes a variety of normal tissue and cell types as well as certain disease states. Many more different biological states, such as rare diseases or particular cell subtypes, exist. It is impractical for a single dedicated experiment to generate a comprehensive expression data set covering all biological conditions, partly owing to cost, but also because some conditions are studied only in specialized laboratories. Even so, we can use computational approaches to integrate the wealth of experiments that already have been performed.

Integration of independent microarray studies is challenging, as microarrays do not measure gene expression in any absolute units. Several studies have integrated single-platform[3] and cross-platform[4–6] data from single-channel oligonucleotide arrays yielding consistent results. It has been generally accepted, however, that only data from the same platform can be reliably integrated on a quantitative level[7]. Integration is also challenging because of the unavoidable complexity of sample descriptions. The Unified Medical Language System has been used to re-annotate free text-based sample descriptions[8]; however, extracting information from published data sets and representing it suitably for statistical analysis is a time-consuming process that is difficult to automate and requires expert curation[9].

We collected over 9,000 raw data files generated on the human gene expression array Affymetrix U133A from the public databases Gene Expression Omnibus[10] and ArrayExpress[11]. After we removed duplicate files and applied strict quality controls (**Supplementary Methods**), data on 5,372 samples from 206 different studies generated in 163 different laboratories remained. Using text mining and curation, we binned the samples in 369 biological groups, each representing a particular cell or tissue type, disease state or cell line (**Supplementary Fig. 1a**). Of these, 96 groups contained at least ten biological replicates. We also introduced 'meta-groups' such as cell lines, neoplasms, non-neoplastic diseases, and normal, as well as groups by tissue of origin (**Supplementary Figs. 1c–e**). The raw data were normalized jointly, producing a gene expression matrix of ~22,000 probe sets (mapping to ~14,000 genes) times 5,372 samples (the complete annotated data set is available from the ArrayExpress repository, accession number E-MTAB-62).

To enable exploration of these data, we have implemented an online query interface (http://www.ebi.ac.uk/gxa/array/U133A). After selecting a particular sample binning (e.g., by tissue of origin), the user can find all genes up- or downregulated in a particular sample class (such as liver). Alternatively, choosing a gene of interest will produce box plots showing the gene's expression across the samples within each of the groups. The coloring of each box plot indicates the outcome of a statistical test for over- or underexpression. Probe set–level queries are also permitted.

As these data were generated in different laboratories, and as laboratory effects are known to be strong[12], it is important to assess the impact of these effects on the analysis. Most laboratories predominantly work with particular types of samples, which makes the lab effects hard to assess. Even so,

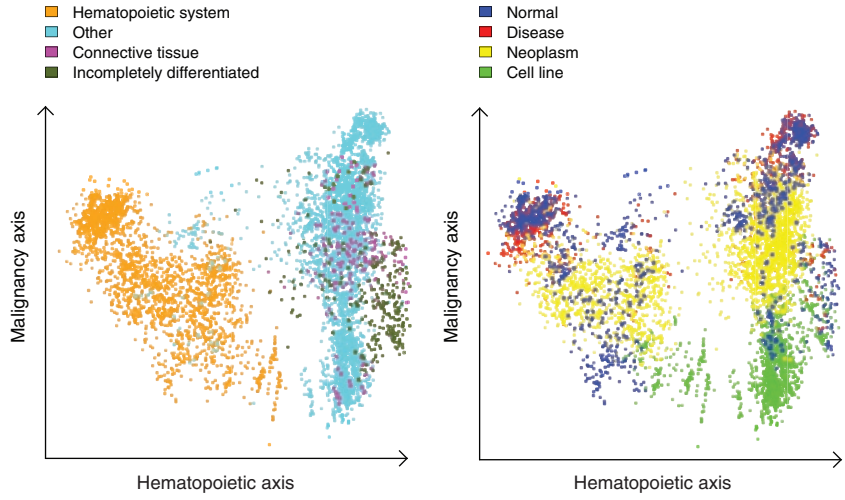


Figure 1 Principal component analysis. Each dot represents one of the 5,372 samples in a multidimensional gene expression space projected on the principal plane formed by the first (hematopoietic) and second (malignancy) principal axes. The dots are colored semitransparently according to the biological group the sample belongs to. (**a**) The first principal component separates hematopoietic system–derived samples from the rest of the samples, with connective tissues and incompletely differentiated cell–based samples forming a relatively compact group on the right. The cyan dots among the blood samples on the right side represent samples from bronchoalveolar lavage cells (a possible sample contamination with blood) and kidney. The dark green dots at the center include embryonic stem cells. (**b**) The second principal axis predominantly arranges cell line samples at the bottom, neoplasm samples in the middle and a mixture of nonneoplastic disease and normal samples at the top.

51 of the 96 larger biological groups (with ten replicates or more) contain assays from at least two different laboratories. In total, 100 different laboratories contributed 3,133 samples to these multi-laboratory biological groups. For each of these biological groups, we computed the average similarity between the assays from different laboratories within the same group. We also computed the average similarity between assays from the same laboratory, but representing different biological groups. The comparison of the two similarity distributions showed that the biological effects were significantly ($P < 2.2^{-16}$) stronger than the laboratory effects (**Supplementary Fig. 2**). For sample classes to which only one laboratory contributed, we cannot distinguish directly between the laboratory and biological effects. However, we can analyze our data from a biological perspective and compare the results to existing knowledge.

We applied principal component analysis (PCA) to the expression matrix, and produced visualizations in which each sample was represented by a point in the plane formed by two principal axes, and colors were assigned to each point according to the biological class (**Fig. 1** and **Supplementary Fig. 3a–e**). We found that the first three principal components have biological interpretations; we named them the hematopoietic, malignancy and neurological axes. Three groups—hematopoietic system, solid tissues and a mixture of incompletely differentiated cell types and connective tissues—were consecutively arranged on the hematopoietic axis. The malignancy axis differentiates three other groups: cell lines, neoplasms and a mixture of normal tissues and non-neoplastic disease tissues. The neurological axis separates nervous system from other samples. The fourth principal component correlates with an array quality metric RLE (relative log expression). The first three principal components explain ~37% of variability in the data (**Supplementary Fig. 3f**). Note that the full expression space consists of thousands of dimensions.

We also used hierarchical clustering to investigate the expression space from a different perspective. We first clustered the 96 larger biological groups (with ≥10 replicates), representing each group by its mean expression profile. Six major clusters emerged: (i) cell lines derived from solid tissues, (ii) incompletely differentiated cell types and connective tissues, (iii) solid normal and neoplastic tissues, (iv) hematopoietic system, (v) brain, and (vi) muscle and heart (**Supplementary Fig. 4a**). This clustering is robust: we obtained similar results when samples from different laboratories were kept in separate groups (**Supplementary Fig. 4b**) and by clustering all 369 sample groups (**Supplementary Fig. 4c**). To see how each of the 273 smaller groups relates to the six original clusters, we computed the pairwise distances between the members of the 96 and 273 groups and applied hierarchical clustering (**Supplementary Fig. 4d**). The smaller-group clusters correspond well to the six original clusters, although an additional small cluster of liver and small-intestine samples emerged. This analysis is driven by the original clustering; nevertheless, if there were new major expression pattern groups, we would expect to observe them. We conclude that the large-scale structure of our data can be explained by six major sample expression profile groups, corresponding to transcriptional states, and some smaller outliers.

Various observations can be made by examining the sample annotations in more detail. For instance, skeletal and heart muscle cluster together, whereas smooth muscle belongs to the incompletely differentiated cell type cluster, which is dominated by fibroblasts. This cluster includes bone-marrow mesenchymal stem cells, but not the hematopoietic bone-marrow stem cells, which are located in the hematopoietic cluster together with other blood-cell precursors. The embryonic stem cell line (HES2; ref. 13) does not belong to the cluster of incompletely differentiated cell types; its expression profile is similar to those of both fibroblasts and neoplastic cell lines.

Next, we studied which genes are expressed in various biological conditions. We applied hierarchical clustering to gene expression profiles across the 96 larger groups, representing the expression of a gene in each group by its mean. We visualized the 1,000 most variable probe sets mapping to 907 different genes and visually identified 50 gene clusters (**Supplementary Fig. 5a**). As our data set represents a wide range of biological conditions, we can study the overall variability of gene expression. For the majority of genes, the normalized signal is largely constant across the 5,372 samples; there are only 1,034 probe sets with a standard deviation >2 (**Supplementary Fig. 6a,b**). The sample clustering obtained using only the 350 most variable probe sets produced similar results to that based on all data and is retained to some extent even when only the 30 most variable probe sets are used (**Supplementary Figs. 4e** and **5b**). Although it is not surprising that only a small number of genes are needed to define six transcriptional states, it is worth noting that the highest expression variance can identify these genes.

To identify genes differentially expressed in specific biological groups, we performed one-way analysis of variance (**Supplementary Methods**). For instance, we found 243 genes differentially expressed in 567 samples grouped under 'leukemia'. Many of these are known to be implicated in leukemia (for example, *BCR*, *ETV6*, *FLT3*, *HOXA9*, *MYST3*, *PRDM2*, *RUNX1* and *TAL1*), and we confirmed many others through literature searches. Similarly, 1,217 genes are differentially expressed in all cell lines: the upregulated genes are most over-represented in gene ontology categories related to M phase, cell division, mitosis, cell cycle and primary metabolic processes, and downregulated genes are most over-represented in immune and defense response.

Our study demonstrates that analysis of a large microarray data set compiled from many laboratories can reveal the overall structure of gene expression space, which could not be observed in any of the contributing studies individually. A particularly important finding is that solid-tissue cell lines form a distinct group, clustering with each other rather than with their respective tissues of origin (**Supplementary Figs. 4a,i**). Moreover, they show high similarity to blood cell lines. An exception to this rule is incompletely differentiated cell types, for which cell lines cluster with the primary cells. Note that on the PCA's malignancy axis, neoplasm samples are located between the cell line and the normal and non-neoplastic disease samples, characterizing neoplasm as an intermediate state between normal samples and immortalized cell lines.

When interpreting these results, several limitations concerning the data set must be taken into account. First, there may be gaps in our data; for instance, there are few normal solid-tissue samples besides muscle, heart and brain. More data may reveal other major transcriptional classes. Second, it is possible that the laboratory effects are too strong to achieve resolution beyond the six major classes. Although the PCA shows samples from more specific groups (such as leukemia) located together (**Supplementary Fig. 3c**), and supervised analysis reveals that genes specific to such sample classes are often known to be involved in the relevant biological conditions, the results of hierarchical clustering did not conclusively reveal finer structures.

To summarize, we have constructed a global map of human gene expression from a large microarray data set. Our analysis reveals six major 'continents' on the map. We acknowledge that there may be more continents that we were not able to find owing to incompleteness of the data, and it is to be expected that finer structures exist within the six we found.

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

ACKNOWLEDGMENTS

We thank P. Adler, J. Bahler, E. Birney, D. Brazma, I. Dunham, N. Gehlenborg, F. Holstege, S. Kaski, M. Krestyaninova, J. Rung, G. Rustici, T. Schlitt and H. Zang-Bradley for helpful comments and discussion. This research was partly funded by FELICS (021902) and EMERALD (LSHG-CT-2006-037689) grants from the European Commission, a MAGE grant from the National Human Genome Research Institute and National Institutes of Biomedical Imaging and Bioengineering from the National Institute of Health (1 P41 HG003619-01), an ALGODAN CoE grant of the Academy of Finland, and grant no. 40274/06 from Finnish Funding agency for Technology and Innovation (TEKES).

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Margus Lukk[1], Misha Kapushesky[1], Janne Nikkilä[2], Helen Parkinson[1], Angela Goncalves[1], Wolfgang Huber[1], Esko Ukkonen[3] & Alvis Brazma[1,4]

[1]European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK. [2]Faculty of Veterinary Medicine and [3]Department of Computer Science, University of Helsinki, Helsinki, Finland.
e-mail: brazma@ebi.ac.uk

1. Su, A.I. *et al. Proc. Natl. Acad. Sci. USA* **99**, 4465–4470 (2002).
2. Su, A.I. *et al. Proc. Natl. Acad. Sci. USA* **101**, 6062–6067 (2004).
3. Day, A., Carlson, M.R., Dong, J., O'Connor, B.D. & Nelson, S.F. *Genome Biol.* **8**, R112 (2007).
4. Kilpinen, S. *et al. Genome Biol.* **9**, R139 (2008).
5. Kapushesky, M. *et al. Nucleic Acids Res.* **38**, D690–D698 (2010).
6. Morgan, A.A., Dudley, J.T., Deshpande, T. & Butte, A.J. *Physiol. Genomics* **40**, 128–140 (2009).
7. Shi, L. *et al. Nat. Biotechnol.* **24**, 1151–1161 (2006).
8. Butte, A.J. & Kohane, I.S. *Nat. Biotechnol.* **24**, 55–62 (2006).
9. Malone, J. *et al. Bioinformatics*, published online 3 March 2010, doi:10.1093/bioinformatics/btq099 (2010).
10. Barrett, T. *et al. Nucleic Acids Res.* **37**, D885–D890 (2009).
11. Parkinson, H. *et al. Nucleic Acids Res.* **37**, D868–D872 (2009).
12. Zilliox, M.J. & Irizarry, R.A. *Nat. Methods* **4**, 911–913 (2007).
13. Hirst, C.E. *et al. Dev. Biol.* **293**, 90–103 (2006).

Empirical analysis of major stem cell patent cases: the role of universities

Ann E Mills & Patti M Tereskerz

A study of stem cell patent litigation sheds light on the current patent reform debate.

It has been theorized[1,2] that litigation is an important contributor to the concept of the anti-commons, which occurs when multiple owners hold the right to exclude each other from a scarce resource, so that no one holds an effective right of entry and underuse of the resource results[3]. This may occur with royalty stacking, where an inventor must obtain multiple licenses to commercialize a product. If this activity is leading to an anti-commons, then, is the multitude of patents that has been granted actually inhibiting, rather than facilitating, the transfer of technology[3]?

Patent litigation and the strategies firms follow to protect themselves from the risk of litigation (e.g., defensive patenting by enlarging a firm's portfolio of patents to influence settlement terms or foregoing otherwise valuable research because of the risk of litigation)[4] generate costs that may divert resources away from innovative activities, or make subsequent commercialization no longer feasible or more costly. Thus, litigation and strategies to avoid litigation can be seen as contributing to an anti-commons.

The perception that a large amount of litigation is occurring is one point advocates use to justify patent reform. For example, the 110th Senate Judiciary Committee's Report on the Patent Reform Act of 2007 warns that excessive litigation creates costs for industry participants[5]. This notion derives from reports by the Federal Trade Commission (FTC) and the National Research Council (NRC)[6–8] that warn of dire consequences if industry is unable to innovate and successfully commercialize new products. Moreover, the risk of being sued for patent violation has now become a major concern to universities since *Madey v. Duke University*[9]. John Madey was a laboratory director at Duke University who received two patents related to free electron laser technology. After internal disagreements, Duke removed him as a director but continued to use his patents. Madey sued, claiming patent infringement. The district court dismissed his claim based on the common law experimental-use doctrine[10] or the idea that infringement is acceptable when done to satisfy "philosophical" or scientific inquiry. The experimental-use exemption was articulated in 1813 by Judge Story[10], who used the term "philosophical" instead of "scientific" to describe the experimental-use exemption from patent infringement. The essential component of the court's reasoning was that those skilled in such "useful arts" are free to use the knowledge imparted by a patent disclosure for amusement, to satisfy idle curiosity or for strictly philosophical inquiry. On appeal, the US Court of Appeals for the Federal Circuit reversed the decision and remanded the case. Duke's petition to the Supreme Court seeking review of the Federal Circuit's decision was denied[11].

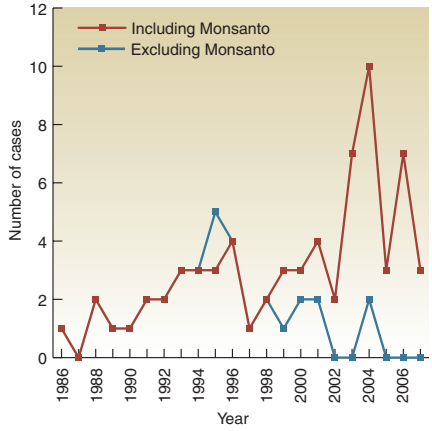


Figure 1 Litigated cases per year.

The Federal Circuit held that the experimental-use exemption for patent infringement does not apply to research that furthers universities' "business objectives," including research and educating and enlightening students and faculty. The court held that "so long as the act is in furtherance of the alleged infringer's legitimate business and is not solely for amusement, to satisfy idle curiosity or for strictly philosophical inquiry, the act does not qualify for the very narrow and strictly limited experimental-use defense"[9]. Moreover, the profit or nonprofit status of the user is not relevant. Thus, the court found that the experimental-use exception should not insulate commercial research from claims of patent infringement. This applied to Duke University, which, as Judge Gajarsa noted, was "not shy in pursuing an aggressive patent licensing program from which it derives a not insubstantial revenue stream"[9].

The FTC and NRC reports question whether the absence of an observed anti-commons is associated with researchers' lack of knowledge about the *Madey* decision[6–8]. As more researchers and their institutions become more knowledgeable about the potential for patent violations following *Madey*, they, as well as private industry, may become more litigious in protecting their intellectual property rights, allowing an anti-commons to emerge.

We studied stem cell patents to determine the rate of litigation of such patents, to describe that litigation and to determine if it has, in fact, been increasing since *Madey*. We chose to study stem cell patent litigation

Ann E. Mills and Patti M. Tereskerz are at the University of Virginia Center for Biomedical Ethics and Humanities, Program in Ethics and Policy in Health Systems, Charlottesville, Virginia, USA.
e-mail: amh2r@virginia.edu

Box 1 Study description

Bergman and Graff identified stem cell patents by using the IPC code in which the technical fields of a patent's primary "inventive steps" are reported[13]. This allowed them to identify the areas where most stem cell patenting has occurred. According to these investigators, the three dominant areas in which stem cell patenting occur are within the IPC codes C07K 14/ (peptides having more than 20 amino acids), C12N 5/ (undifferentiated human, animal or plant cells, e.g., cell lines; tissues; cultivation or maintenance thereof; culture media) and C12 15/ (mutation or genetic engineering; DNA or RNA concerning genetic engineering vectors, their isolation, preparation or purification) (http://www.wipo.int/classifications/ipc/ipc8/?lang=en).

We completed this study in two phases. During the week of June 30 to July 4, 2008, we accessed the USPTO database (http://patft.uspto.gov/) using these codes for phase 1. In phase 2, using these codes, the USPTO supplied us with a list of patents issued between January 1, 2006 and February 9, 2010 (this list is available on request). The reasons for completing the study in two phases include updating this study at the editor's request and the fact that an error on the full-text USPTO database prevents us from searching on these codes after January 1, 2006. This error has been reported to and acknowledged by the USPTO. In both phases, we took a 30% random sample of patents in each area identified by Bergman and Graff to be included in our study.

For the first phase, we cross-referenced the patents we identified with the LitAlert database, which contains records for patent lawsuits filed in the 94 US District Courts and reported to the commissioner of the USPTO. In addition, records for thousands of lawsuits filed since the early 1970s that have not been reported in the USPTO *Official Gazette for Patents*[18] are included. We searched the LitAlert database for the patent numbers contained in each of our three data sets. During the second phase of the study, we followed the same method beginning and ending our cross-reference to LitAlert on February 15, 2010.

Many cases involve more than one patent. LitAlert identifies a primary patent and lists any other patent that might be part of the litigated case. To avoid multiple counting of lawsuits, in both phases, we selected only those cases where the primary patent number was listed in our data sets. Once we identified the appropriate cases, we cross-referenced the patent numbers involved in the cases with the USPTO database to obtain the issue dates of the patents.

for two reasons. First, the area of stem cell research and technology is now once again at the forefront of the nation's research agenda, since President Barack Obama's executive order removed some limitations from stem cell research[12]. The research has potential to yield new treatments and cures for debilitating and fatal diseases so it is vital that the US patent system encourage innovation in this area to maintain global competitiveness.

Second, concerns about an emerging anti-commons are particularly acute in stem cell research, where there are few or no alternatives to patented technologies that enable embryonic or adult stem cells to be directed into specific cell lineages. The interdependence of the technologies involved with these stem cell patents and the fragmentation of their ownership across many organizations could make the task of coordinating access to key technologies an intensive as well as costly process, and so create an anti-commons in the field[13].

Results

We built on the work of Bergman and Graff[13] who identified three dominant areas where most stem cell patenting has occurred. They are within three IPC codes: C07K 14/ (peptides having more than 20 amino acids), C12N 5/ (undifferentiated human, animal or plant cells, e.g., cell lines; tissues; cultivation or maintenance thereof; culture media) and C12 15/ (mutation or genetic engineering; DNA or RNA concerning genetic engineering vectors, their isolation, preparation or purification) (See **Box 1**).

In phase 1 of our study, we identified 67 cases involving 23 different patents, meaning many patents were litigated multiple times. Of note is one patent[14], owned by Monsanto, a multinational agricultural biotech company, which was litigated 32 times. In phase 2, we identified five cases involving five patents.

In phase 1, there were no lawsuits in the category encompassing "peptides having more than 20 amino acids" (C07K 14/). There were significantly ($P<0.001$) more law suits for stem cell patents involving "mutation or genetic engineering; DNA or RNA concerning genetic engineering vectors, their isolation, preparation or purification" (C12 15/) compared to stem cell patents for "undifferentiated human, animal or plant cells, e.g., cell lines; tissues; cultivation or maintenance thereof; culture media" (C12N 5/). However, when you compare the rate of litigation by patent, taking into account that some patents were litigated multiple times, there is no significant difference between these two categories (**Table 1**).

In phase 2 of our study, there were three lawsuits in the category encompassing peptides having more than 20 amino acids (C07K 14/) and one lawsuit in each of the other categories. There were no significant differences in litigation between the categories studied (**Table 2**).

We identified the plaintiffs in each lawsuit. Some lawsuits involved more than one plaintiff. We included all plaintiffs we were able to identify. In phase 1, 7 out of 67 (10%) cases involved university plaintiffs. Of the 28 plaintiffs we were able to identify, 5 (18%) were university plaintiffs (**Supplementary Table 1**). In phase 2, we did not identify any university plaintiffs (**Supplementary Table 2**).

Similarly, a case can involve more than one defendant. In phase 1, we identified 89 defendants, most of whom were involved in only one case. Five of the defendants were involved in two cases. Only one defendant was a university plaintiff (**Supplementary Table 3**). In phase 2, we identified three defendants. One of the defendants, a representative of the US Patent and Trademark Office (USPTO), was involved in three cases (**Supplementary Table 4**).

In phase 1, we calculated the difference in months between the time a patent issues and a suit is filed. The difference in months between the time a patent issues and a suit is filed is shown in **Supplementary Table 5**. The average time difference by category for 64 cases is shown in **Supplementary Table 6**. (We eliminated three cases for this calculation because the filing dates given by LitAlert (the patent and trademark litigation alert database published by Westlaw) are the same as issue dates. This is highly improbable and can probably be attributed to input error. One case in category C12N 5/ was eliminated and two cases in category C12 15/ were eliminated.) Because Monsanto seems to be following a deliberate policy of suing over older patents[14,15] (the first, litigated 32 times, issued in 1994; and the second, litigated 3 times, issued in 1990) (**Supplementary Table 1**), we also calculated an average time for those cases in which Monsanto was not a plaintiff (**Supplementary Table 6**).

Similarly, in phase 2, we calculated the difference in months between the time a patent issues and litigation. However, we are doubtful that these data are meaningful as we identified so few cases associated with these later-issuing patents. However, the fact that we identified so few cases could validate phase 1's results that, on average, it takes over 62 months (excluding lawsuits involving Monsanto as plaintiff) for a suit to be filed after a patent issues (**Supplementary Table 7**).

In phase 1, we calculated the number of cases litigated per year including and excluding Monsanto. Again, we eliminated the three cases which seemed to have incorrect filing dates. (We cannot obtain a rate of litigation for the total number of patents for all IPC codes studied, nor can we get a rate of litigation per patent because the total number of patents identified is inflated. For example, a patent might have the code C07K 14/ and the C12 15/ code and so the total would contain duplicate patents.) The results are shown in **Figure 1**. In phase 2, because there were so few cases, we did not calculate number of cases litigated per year.

Table 2 Study phase 2: litigation of stem cell patents from January 1, 2006 to February 9, 2010

IPC code	C07K 14/	C12N 5/	C12 15/	*P* value
Number of patents identified on USPTO database	1,683	1,096	2,118	
N (total patents in this study)	506	329	635	
Number of litigated cases	3	1	1	0.45 (Chi squared or $P = -0.53$ by Fisher's Exact Test)
Rate of litigated cases	0.006	0.003	0.002	
Number of patents litigated	3	1	1	
Rate of number of patents in study	0.006	0.003	0.002	0.45 (Chi squared or $P = -0.53$ by Fisher's Exact Test)

Discussion

For the stem cell patents in both phase 1 and 2, the rate of litigation to issued patents in the categories studied is extremely small, which calls into question the claim that a large amount of litigation is causing an anti-commons. However, when a patent is litigated, often it is litigated more than once, meaning that relatively few stem cell patents account for a large portion of litigation for the categories of patents we studied.

Contrary to the idea that industry is responsible for almost all patent litigation, it appears that for stem cell patent litigation, universities have already assumed the role of plaintiff. This supports the Federal Circuit's finding in *Madey* that universities have their own business objectives and therefore should not be a protected class that benefits from a research exemption when the research furthers their business interests. Further evidence that universities seek to protect their intellectual property interests is that in phase 1, excluding the Monsanto patents, of the seven patents that were litigated more than once, three of these patents were associated with universities. Of these 12 cases, 7 cases involved universities (**Supplementary Table 1**).

Only one of the defendants we identified was a university. *Madey* was decided in October 2002. Because 46% of the litigation we studied in phase 1 was after *Madey*, and the average time between patent issue and institution of a lawsuit is 62.52 months (using the average time, excluding Monsanto), it appears that the post-*Madey* fears of opening a floodgate of litigation against universities is not being realized at the moment, at least for these stem cell patent categories. The one case we found with a university being a defendant was filed in 1999, which predates *Madey* by several years. These results are supported by phase 2 of the study in which we did not find a university named as defendant.

These findings are consistent with two surveys Walsh and colleagues undertook[8,16]. In the first survey, they conducted 70 interviews with intellectual property attorneys, business managers and scientists from ten pharmaceutical and biotech companies, as well as others, including university researchers and technology transfer officers, patent lawyers and trade association personnel[14]. They found that university researchers, to the extent that they are doing noncommercial work, are left alone. In fact, Walsh and colleagues found that industry welcomes research like this because it helps develop more knowledge about the patented technology. Moreover, industry respondents suggested that companies, because of the negative publicity that usually attends assertive action against a university, believe it is not worth bringing such suits.

The National Research Council arranged with Walsh and colleagues to undertake a second, more extensive survey[8] than the first. A sample of 1,124 persons included, among others, investigators in universities, government laboratories and other nonprofit institutions, 563 industry scientists, and 299 researchers working on a signaling protein. The findings of the second study are similar to those of the first in terms of technology access issues. For example, in the sample above, "unreasonable terms for obtaining research inputs" was cited by 10% of survey respondents and "too many patents covering needed research inputs" was cited by only 3% of survey respondents as a reason for project abandonment[8].

Table 1 Study phase 1: litigation of stem cell patents before January 2006

IPC code	C07K 14/	C12N 5/	C12 15/	*P* value
Number of patents identified on USPTO database (June 30–July 4, 2008)	2,019	3,878	6,980	
N (total patents in this study)	605	1,163	2,094	
Number of litigated cases	0	9	58	
Rate of litigated cases	0	0.00773861	0.02769819	$P < 0.001$ (likelihood ratio test for comparing Poisson counts)
Number of patents litigated	0	6	17	
Rate of number of patents in study	0	0.00515907	0.00811843	Not significant $P = 0.068$ (Chi square)

Study period June 30–July 4, 2008.

Limitations and conclusion

There are limitations to this study. First, we studied only one type of patent—stem cell patents—in three categories. Therefore, the results of this small study cannot be extrapolated to other industries or areas. Secondly, we only studied litigation, which is theorized to be but one contributor to the anti-commons, albeit thought to be an important one.

There are also the study's technical limitations. Federal law requires clerks of the courts of the United States to notify the USPTO

within one month after the filing of a suit[15]. LitAlert obtains information for its database from the USPTO. But not all courts may be diligent about this requirement. Thus, if filed suits are not reported, they will not be included in LitAlert's database.

We find no evidence that *Madey* is unleashing a floodgate of litigation in the areas we studied, despite the worry voiced by the three reports cited above. We also find that for the stem cell patents studied, universities are proactive (and have been before *Madey*) in protecting their intellectual property, supporting the *Madey* contention that the boundary between industry and academia is blurring.

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

ACKNOWLEDGMENTS

The authors would like to acknowledge M. Conway who performed the statistical analysis.

COMPETING FINANCIAL INTERESTS

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

1. Allison, J.R. *et al. Geo. L.J.* **92**, 435–476 (2004).
2. Parchomovsky, G. & Wagner, R.P. *U. Pa. L. Rev.* **154**, 1–77 (2005).
3. Heller, M. & Eisenberg. R.S. *Sci.* **280**, 698–701 (1998).
4. Azher. A. I. *U. Pa. J. Int'l. Econ. L.* **25**, 383–421 (2004).
5. 110th Senate Judiciary Committee's Report on the Patent Reform Act of 2007. <http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=110_cong_reports&docid=f:sr259.pdf>
6. Federal Trade Commission. *To Promote Innovation: The Proper Balance of Competition and Patent Law and Policy* (FTC, 2003).
7. Merrill, S. *et al.* (eds.) A *Patent System for the 21st Century* (National Academies Press, Washington, DC, 2004).
8. National Research Council. Committee on Intellectual Property Rights in Genomic and Protein Research and Innovation. *Reaping the Benefits of Genomic and Protemic Research: Intellectual Property Rights, Innovation, and Public Health* (National Academies Press, 2006).
9. *Madey v. Duke University*, 307 F.3d 1351 (N.C. 2002).
10. *Whittemore v. Cutter*, 29 F. Cas. 1120 1121 (C.C.D. Mass 1813).
11. Cert. denied 123 S. Ct. 2639 (2003).
12. Obama, B. Removing barriers to responsible scientific research involving human stem cells (The White House, 9 March, 2009) <http://www.whitehouse.gov/the_press_office/Removing-Barriers-to-Responsible-Scientific-Research-Involving-Human-Stem-cells/>
13. Bergman, K. & Graff, G.D. *Nat. Biotechnol.* **25**, 419–424 (2007).
14. US Patent 5,352,605.
15. US Patent 4,940,835.
16. Walsh, J.P., Arora, A. & Cohen, W. in *Patents in the Knowledge-Based Economy*, (eds. Cohen W. & Merrill, S.) 285–340 (National Academies Press, Washington, DC, 2003).
17. 35 USC 290, Notice of Patent Suits.
18. http://www.uspto.gov/news/og/pindex.jsp

Recent patent applications in high-throughput drug screening

Patent number	Description	Assignee	Inventor	Priority application date	Publication date
US 20100056390	A three-dimensional cell culture system comprising a cell of interest and a natural or synthetic hydrogel incorporating the cell of interest; useful in an apparatus for high-throughput drug screening.	Fischbach C	Fischbach C	7/29/2008	3/4/2010
US 20100041901	A new fluorescent compound having a central linker element linking two 6,5-ring systems or two 6,6,5-ring systems; useful as a fluorescent reporter molecule in a biological assay for drug discovery and as a fluorescent marker for protein identification.	Cairns N	Cairns N	8/14/2008	2/18/2010
WO 2010019388, US 20100041046	A method of discretizing and manipulating sample volumes to perform a digital polymerase chain reaction, comprising providing a fluidic lattice with a flow channel and many fluidic harbors.	University of Washington (Seattle)	Chiu DT, Cohen DE, Cohen DW, Jeffries GDM	8/15/2008	2/18/2010, 2/18/2010
US 7659322	A method of forming polyanhydride for controlled drug delivery involving irradiating diacid with microwave radiation in the presence of carboxylic anhydride to acylate the diacid to form a prepolymer, and irradiating the prepolymer with microwave radiation to polymerize and yield the polyanhydride.	Iowa State University Research Foundation (Ames, IA, USA)	Mallapragada SK, Vogel BM	12/16/2004	2/9/2010
US 20100028892	A reporter gene system comprising an exogenous protease with or without an exogenous activator of protease, which cleaves transcription factor protein; useful for monitoring protease activity in living yeast, e.g., *Saccharomyces cerevisiae*.	Burnham Institute for Medical Research (La Jolla, CA, USA)	Cuddy M, Hayashi H, Reed JC	7/16/2008	2/4/2010
WO 2010011939, US 20100022416	An etched plate comprising a substrate, a first layer partially disposed on the substrate and etched features; useful for measuring electrode-induced property in an etched features assay array having >5,000 etched features.	Life BioScience (Albuquerque, NM, USA)	Buckley CT, Flemming JH, Ridgeway B	7/25/2008	1/28/2010, 1/28/2010
WO 2010008519	A marker comprising a binding agent, e.g., biotin, and a chromophore, e.g., fluorescein; useful for investigating analytes *in vitro* and for anatomic, physiologic and biochemical parameters.	Hu H, Li S, Metters A, Wang Q	Hu H, Li S, Metters A, Wang Q	7/14/2008	1/21/2010
WO 2010003908	A biological system for analyses of the inhibitory mode of action exerted by a compound on the ubiquitin proteasome system for cellular protein degradation, comprising expression of heterologous fusion proteins in a eukaryotic host cell.	Esmaeilzadeh S, Jakobsson A, Pouya A	Beskow A, Bjoerk Grimberg K, Bott LC, Dantuma N, Salomons F, Young P	7/8/2008	1/14/2010
US 20090246802, WO 2009124135	A new annexin derivative comprising amino acids conjugated to a polarity-sensitive fluorophore; useful for detecting apoptosis and for monitoring cell health.	University of Southern California (Los Angeles)	Langen R	4/1/2008	10/1/2009, 10/8/2009
WO 2009047760	A new apparatus for measuring protein translation, comprising a cell or subcellular compartment where the cell or subcellular compartment comprises at least one protein synthesis element.	Anima Cell Metrology (Bernardsville, NJ, USA)	Smilansky Z	10/9/2007	4/16/2009

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).

Broad-spectrum defense against plant pathogens

Alexandre Brutus & Sheng Yang He

Transfer of a pattern-recognition immune receptor to a crop confers resistance to several bacterial pathogens.

Like animals, plants must continually fight infection by the many pathogens they encounter. Microbial pathogens account for billions of dollars worth of crop losses annually, and genetic engineering solutions that could defend plants against a broad range of pathogens would have a large impact on agricultural productivity. In this issue, Lacombe *et al.*[1] report a groundbreaking demonstration that heterologous expression of a plant immune receptor can provide broad-spectrum resistance against four agriculturally relevant bacterial pathogens.

Plants use two types of immune receptor to combat microbial pathogens. First, pattern recognition receptors located on the plant cell membrane detect so-called pathogen-associated molecular patterns (PAMPs), molecules such as lipopolysaccharide and flagellin that animals and plants use to recognize groups of pathogens. This results in PAMP-triggered immunity. Second, plants possess disease-resistance proteins (R proteins), usually located inside the plant cell. R proteins detect pathogen effector proteins secreted into the plant cell during infection, leading to effector-triggered immunity. Many of the endpoints in PAMP-triggered and effector-triggered immunity are the same, but the effector-triggered response is stronger and may involve activation of additional defense signaling, including programmed plant cell death.

Effector-triggered immunity can be highly effective against single pathogens and has been exploited in conventional and transgenic breeding programs to enhance resistance to crop diseases. In contrast, the feasibility of using pattern-recognition receptors in crop protection has remained untested, in part because, until recently, the role of PAMP-triggered immunity in plant disease resistance has not been widely recognized[2]. Unlike R proteins, pattern-recognition receptors should be able to confer resistance to a broad range of pathogens. Moreover, PAMPs tend to be more highly conserved and more essential to microbial viability than effector proteins, suggesting that disease-resistance strategies based on pattern recognition receptors would yield more durable resistance under field conditions compared with 'gene-for-gene' approaches involving R proteins.

Most plant pattern-recognition receptors belong to the family of leucine-rich repeat receptor kinases, which consist of an extracellular leucine-rich repeat domain involved in pathogen recognition, a transmembrane domain and a cytoplasmic serine/threonine kinase domain (**Fig. 1**). Binding of these receptors to cognate PAMPs initiates kinase signaling cascades that activate many downstream immune responses, including production of reactive oxygen species, expression of immunity-associated genes, stomatal closure, callose deposition at the cell wall, and disease resistance[2].

Lacombe *et al.*[1] focused on elongation factor Tu receptor (EFR) from *Arabidopsis thaliana*, one of two pattern-recognition receptors that have been particularly well studied (with the other being flagellin-sensitive 2 (FLS2)). EFR recognizes elongation factor Tu (EF-Tu), a widely conserved bacterial PAMP[3,4]. Cytosolic EF-Tu is involved in translation, whereas an extracellular form of the protein seems to help bacteria adhere to host cells[5]. Elf18, an 18-amino acid peptide at the N terminus of EF-Tu, shows a high degree of amino acid

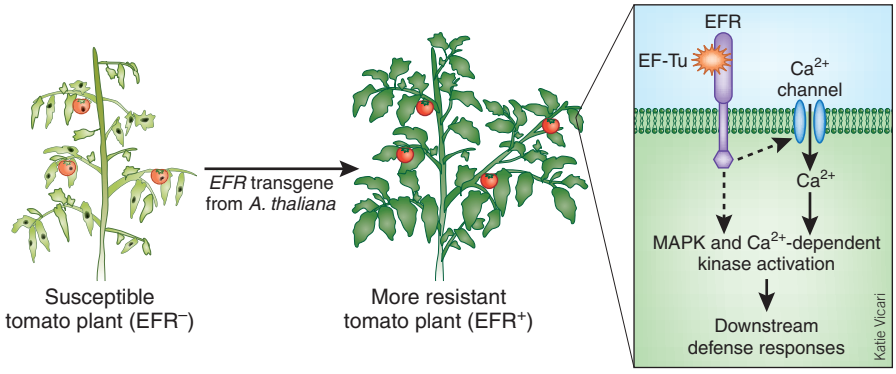


Figure 1 Engineering broad-spectrum resistance in tomato, a member of the Solanaceae, by transferring a pattern-recognition receptor from the wild species *A. thaliana* (family Brassicaceae). Elongation factor Tu receptor (EFR; purple) is absent from the solanaceous species *N. benthamiana* and tomato, and these plants are normally susceptible to infection by *Agrobacterium tumefaciens*, *Pseudomonas syringae*, *Ralstonia solanacearum* and *Xanthomonas perforans*. Transgenic expression of *A. thaliana* EFR increases their resistance to all of these bacterial pathogens, presumably by activating a signaling cascade that confers resistance to a range of bacteria expressing its cognate PAMP, elongation factor Tu (EF-Tu; orange). The successful inter-family transfer of EFR-mediated disease resistance suggests that *N. benthamiana* and tomato contain all components necessary for EFR signaling other than the receptor. MAPK, mitogen-activated protein kinase.

Alexandre Brutus and Sheng Yang He are in the Department of Energy Plant Research Laboratory, Michigan State University, East Lansing, Michigan, USA.
e-mail: hes@msu.edu

conservation among many bacterial species[3] and is used as a surrogate in studies of EFR-triggered defense responses.

Assays of elf18-induced immune responses (e.g., H_2O_2 and ethylene production) have suggested that EFR is expressed in *A. thaliana* and other members of the Brassicaceae but not in members of the Solanaceae[3], such as tobacco, tomato and potato. Heterologous transient expression of *A. thaliana* EFR in *Nicotiana benthamiana*, a close relative of tobacco, confers responsiveness to elf18, as measured by activation of H_2O_2 and ethylene production[4]. Until now, however, it was unclear whether stable transgenic expression of EFR in EFR-lacking crop plants would impart disease resistance, and if so, whether the resistance would be effective against a broad spectrum of bacterial pathogens.

Although much remains to be done, the exciting demonstration of EFR-mediated broad-spectrum disease resistance is likely to stimulate a wave of similar experiments aimed at using immune-enhancing genes to fight plant diseases.

Lacombe *et al.*[1] have answered both questions in the affirmative, showing that EFR transgenic *N. benthamiana* and tomato plants exhibit considerably increased resistance to diverse pathogens from the bacterial genera *Pseudomonas*, *Agrobacterium*, *Xanthomonas* and *Ralstonia*. The resistance appears to depend on EF-Tu recognition, as there is no evidence of activation of nonspecific, constitutive defense responses. Moreover, the resistance is restricted to EF-Tu–expressing bacterial pathogens and is not effective against a fungal pathogen. Importantly, the transgenic plants resemble nontransgenic wild-type plants in size and morphology, with no obvious growth defects, at least in the laboratory setting. The inter-family transfer of EFR-mediated disease resistance suggests that *N. benthamiana* and tomato contain all necessary components for EFR signaling, a finding that contrasts with the difficulty encountered in inter-family transfer of gene-for-gene resistance conferred by *A. thaliana* R proteins.

The efficacy of EFR-mediated immunity compared with R protein–mediated immunity is still unclear. Many laboratory assays have shown that PAMP-triggered immunity has a weaker collective immune output than effector-triggered immunity. Accordingly, Lacombe *et al.*[1] show that in tomato, EFR-mediated resistance to *Xanthomonas perforans* is weaker than resistance mediated by the R protein Bs2. Moreover, it remains to be seen whether pattern-recognition receptor–mediated immunity will indeed be more durable than R protein–mediated immunity. Here, EFR may have an advantage over most R proteins because it recognizes a highly conserved PAMP. R protein–mediated resistance frequently breaks down owing to mutations in the cognate pathogen effector or horizontal acquisition of an immune-suppressing effector. Similarly, the durability of EFR-mediated resistance could be undermined by resistance-evading mutations in PAMPs or by immune-suppressing pathogen effectors, both of which have been reported[6,7]. In addition to providing more data about effectiveness and durability, field trials will better reveal any fitness costs associated with heterologous expression of pattern-recognition receptors.

Too often, promising laboratory demonstrations of engineered disease resistance fall short in field tests. The few success stories include antisense and coat-protein overexpression strategies for controlling plant viral diseases. Although much remains to be done, the exciting demonstration of EFR-mediated broad-spectrum disease resistance is likely to stimulate a wave of similar experiments aimed at using immune-enhancing genes to fight plant diseases. To date, only a handful of pattern-recognition receptors have been identified in plants and, for historical reasons, most of these recognize bacterial pathogens. Of note, the rice R protein Xa21, initially identified through conventional breeding, was recently shown to be a pattern-recognition receptor[8]. With the increasing interest in PAMP-triggered immunity, it is likely that many more pattern-recognition receptors for bacteria and other pathogens will soon be discovered. The pioneering study by Lacombe *et al.*[1] will undoubtedly provide a useful guide in future efforts to engineer pattern recognition receptor–based resistance against a variety of plant diseases caused by bacteria and other prokaryotes, fungi, oomycetes and possibly even nematodes.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

1. Lacombe, S. *et al. Nat. Biotechnol.* **28**, 365–369 (2010).
2. Boller, T. & Felix, G. *Annu. Rev. Plant Biol.* **60**, 379–406 (2009).
3. Kunze, G. *et al. Plant Cell* **16**, 3496–3507 (2004).
4. Zipfel, C. *et al. Cell* **125**, 749–760 (2006).
5. Dallo, S.F., Kannan, T.R., Blaylock, M.W. & Baseman, J.B. *Mol. Microbiol.* **46**, 1041–1051 (2002).
6. Sun, W. *et al. Plant Cell* **18**, 764–779 (2006).
7. Boller, T. & He, S.Y. *Science* **324**, 742–744 (2009).
8. Lee, S.W. *et al. Science* **326**, 850–853 (2009).

Antagonizing metastasis

Michele De Palma & Luigi Naldini

Therapeutic inhibition of a microRNA reduces metastasis formation in a mouse model of breast cancer.

Although metastases are more deadly than primary tumors, the development of anti-metastatic therapies has been hampered by a limited understanding of the underlying biology and a resulting lack of suitable drug targets[1]. In this issue, Weinberg and colleagues[2] demonstrate the feasibility and therapeutic efficacy of targeting a pro-metastatic microRNA (miRNA) with antagomirs. Although the therapy did not affect the growth of the primary tumor, it effectively inhibited metastatic spread to the lung in a mouse model of breast cancer.

miRNAs are an abundant class of small non-coding RNAs that regulate gene expression post-transcriptionally by inhibiting mRNA translation and/or inducing its degradation[3]. miRNAs recognize their targets in a sequence-specific manner and, whereas the downregulation of any individual target is minor, each miRNA targets multiple genes, often linked in a biological pathway, and can thus have a profound impact on cell fate and behavior.

A strong connection has emerged between miRNA and cancer[4]. Indeed, many miRNAs

San Raffaele Telethon Institute for Gene Therapy and Vita Salute San Raffaele University, Milan, Italy
e-mail: naldini.luigi@hsr.it

conservation among many bacterial species[3] and is used as a surrogate in studies of EFR-triggered defense responses.

Assays of elf18-induced immune responses (e.g., H_2O_2 and ethylene production) have suggested that EFR is expressed in *A. thaliana* and other members of the Brassicaceae but not in members of the Solanaceae[3], such as tobacco, tomato and potato. Heterologous transient expression of *A. thaliana* EFR in *Nicotiana benthamiana*, a close relative of tobacco, confers responsiveness to elf18, as measured by activation of H_2O_2 and ethylene production[4]. Until now, however, it was unclear whether stable transgenic expression of EFR in EFR-lacking crop plants would impart disease resistance, and if so, whether the resistance would be effective against a broad spectrum of bacterial pathogens.

Although much remains to be done, the exciting demonstration of EFR-mediated broad-spectrum disease resistance is likely to stimulate a wave of similar experiments aimed at using immune-enhancing genes to fight plant diseases.

Lacombe *et al.*[1] have answered both questions in the affirmative, showing that EFR transgenic *N. benthamiana* and tomato plants exhibit considerably increased resistance to diverse pathogens from the bacterial genera *Pseudomonas*, *Agrobacterium*, *Xanthomonas* and *Ralstonia*. The resistance appears to depend on EF-Tu recognition, as there is no evidence of activation of nonspecific, constitutive defense responses. Moreover, the resistance is restricted to EF-Tu–expressing bacterial pathogens and is not effective against a fungal pathogen. Importantly, the transgenic plants resemble nontransgenic wild-type plants in size and morphology, with no obvious growth defects, at least in the laboratory setting. The inter-family transfer of EFR-mediated disease resistance suggests that *N. benthamiana* and tomato contain all necessary components for EFR signaling, a finding that contrasts with the difficulty encountered in inter-family transfer of gene-for-gene resistance conferred by *A. thaliana* R proteins.

The efficacy of EFR-mediated immunity compared with R protein–mediated immunity is still unclear. Many laboratory assays have shown that PAMP-triggered immunity has a weaker collective immune output than effector-triggered immunity. Accordingly, Lacombe *et al.*[1] show that in tomato, EFR-mediated resistance to *Xanthomonas perforans* is weaker than resistance mediated by the R protein Bs2. Moreover, it remains to be seen whether pattern-recognition receptor–mediated immunity will indeed be more durable than R protein–mediated immunity. Here, EFR may have an advantage over most R proteins because it recognizes a highly conserved PAMP. R protein–mediated resistance frequently breaks down owing to mutations in the cognate pathogen effector or horizontal acquisition of an immune-suppressing effector. Similarly, the durability of EFR-mediated resistance could be undermined by resistance-evading mutations in PAMPs or by immune-suppressing pathogen effectors, both of which have been reported[6,7]. In addition to providing more data about effectiveness and durability, field trials will better reveal any fitness costs associated with heterologous expression of pattern-recognition receptors.

Too often, promising laboratory demonstrations of engineered disease resistance fall short in field tests. The few success stories include antisense and coat-protein overexpression strategies for controlling plant viral diseases. Although much remains to be done, the exciting demonstration of EFR-mediated broad-spectrum disease resistance is likely to stimulate a wave of similar experiments aimed at using immune-enhancing genes to fight plant diseases. To date, only a handful of pattern-recognition receptors have been identified in plants and, for historical reasons, most of these recognize bacterial pathogens. Of note, the rice R protein Xa21, initially identified through conventional breeding, was recently shown to be a pattern-recognition receptor[8]. With the increasing interest in PAMP-triggered immunity, it is likely that many more pattern-recognition receptors for bacteria and other pathogens will soon be discovered. The pioneering study by Lacombe *et al.*[1] will undoubtedly provide a useful guide in future efforts to engineer pattern recognition receptor–based resistance against a variety of plant diseases caused by bacteria and other prokaryotes, fungi, oomycetes and possibly even nematodes.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

1. Lacombe, S. *et al. Nat. Biotechnol.* **28**, 365–369 (2010).
2. Boller, T. & Felix, G. *Annu. Rev. Plant Biol.* **60**, 379–406 (2009).
3. Kunze, G. *et al. Plant Cell* **16**, 3496–3507 (2004).
4. Zipfel, C. *et al. Cell* **125**, 749–760 (2006).
5. Dallo, S.F., Kannan, T.R., Blaylock, M.W. & Baseman, J.B. *Mol. Microbiol.* **46**, 1041–1051 (2002).
6. Sun, W. *et al. Plant Cell* **18**, 764–779 (2006).
7. Boller, T. & He, S.Y. *Science* **324**, 742–744 (2009).
8. Lee, S.W. *et al. Science* **326**, 850–853 (2009).

Antagonizing metastasis

Michele De Palma & Luigi Naldini

Therapeutic inhibition of a microRNA reduces metastasis formation in a mouse model of breast cancer.

Although metastases are more deadly than primary tumors, the development of anti-metastatic therapies has been hampered by a limited understanding of the underlying biology and a resulting lack of suitable drug targets[1]. In this issue, Weinberg and colleagues[2] demonstrate the feasibility and therapeutic efficacy of targeting a pro-metastatic microRNA (miRNA) with antagomirs. Although the therapy did not affect the growth of the primary tumor, it effectively inhibited metastatic spread to the lung in a mouse model of breast cancer.

miRNAs are an abundant class of small non-coding RNAs that regulate gene expression post-transcriptionally by inhibiting mRNA translation and/or inducing its degradation[3]. miRNAs recognize their targets in a sequence-specific manner and, whereas the downregulation of any individual target is minor, each miRNA targets multiple genes, often linked in a biological pathway, and can thus have a profound impact on cell fate and behavior.

A strong connection has emerged between miRNA and cancer[4]. Indeed, many miRNAs

San Raffaele Telethon Institute for Gene Therapy and Vita Salute San Raffaele University, Milan, Italy
e-mail: naldini.luigi@hsr.it

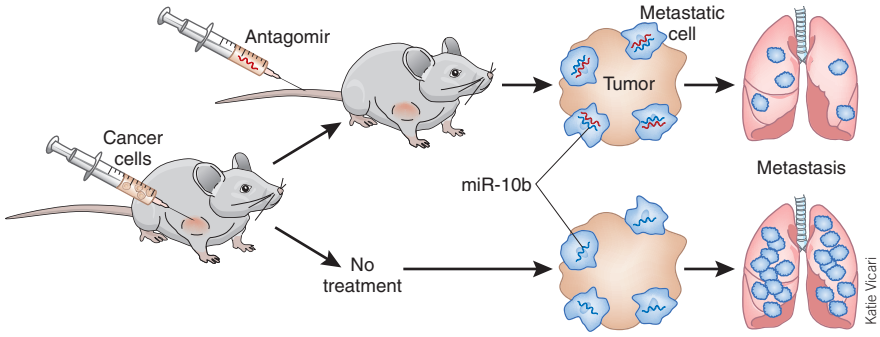


Figure 1 Antagomir treatment to prevent metastatic spread of breast cancer in mice. A tumor is established by implanting highly metastatic tumor cells into the mammary fat pad. Metastatic tumor cells express high levels of the miRNA miR-10b and untreated mice rapidly develop many lung metastases. In contrast, systemic treatment with an antagomir that inhibits miR-10b significantly reduces the number of metastases despite not affecting the growth of the primary tumor.

have been found to be either up- or down-regulated in tumors compared to the tissue of origin. More recently, miRNAs have been implicated in the regulation of metastasis[5], the process by which tumor cells colonize distant organs. Previous work by the Weinberg laboratory[6,7] showed that Twist—a master embryonic transcription factor regulating epithelial-mesenchymal transition—can trigger invasion and metastasis, and that the miRNA miR-10b is a functionally important transcriptional target of Twist that regulates a molecular cascade culminating in the execution of the metastatic program. The analysis of a small set of human breast cancers also suggested that miR-10b was more highly expressed in the metastatic as compared to the nonmetastatic tumors[7]. Although studies using more patients are needed to establish the clinical relevance of these findings[8], miR-10b has been associated with high-grade malignancy in several types of cancer[9]. Thus, antagonizing miR-10b activity may represent a promising therapeutic strategy to inhibit the invasive and metastatic behavior of cancer cells.

To silence the miRNA of interest, the authors used an antagomir—an antisense RNA oligonucleotide chemically modified to improve stability and facilitate cell penetration[10] (**Fig. 1**). Upon systemic administration, antagomirs have been shown to effectively inhibit the activity of endogenous miRNAs in a sequence-specific manner and in many cells types of the body. After multiple injections of a miR-10b antagomir into tumor-bearing mice, Weinberg and colleagues[2] observed markedly reduced miR-10b levels in the tumor. More importantly, one of its validated targets, Hoxd10, was upregulated, demonstrating suppression of miR-10b regulatory activity in the tumor cells. MiR-10b silencing did not inhibit the growth of the primary tumor, but drastically decreased the number of pulmonary metastases compared to mice treated with a control antagomir with a scrambled sequence. Similarly, antagomir treatment inhibited cell motility and invasion, but not proliferation *in vitro*, supporting the concept that miR-10b may specifically regulate the ability of tumor cells to disseminate to distant organs.

> "In the future, one may conceive testing miR-10b antagomirs for the prophylactic treatment of tumors that have not yet metastasized, or to counteract progression to metastasis possibly induced by adaptive resistance to antiangiogenic therapy."

Although overexpressed in the tumor cells, mir-10b is also active in many normal tissues. Because systemically administered antagomirs can penetrate most cells, Weinberg and colleagues[2] needed to exclude the possibility that the antagomir prevented metastasis by silencing miR-10b in normal cells that supported the metastatic process. Host tissues distant from the site of the primary tumor can in fact provide the soil for the metastatic spreading of tumor cells, and several types of reciprocal tumor-host interactions are relevant in the process[1]. To address this issue, the authors used a genetic strategy based on overexpressing an artificial transcript containing the miRNA target sequences. Accumulation of this transcript in the cell cytoplasm acts as a decoy or 'sponge' for the cognate miRNA, competitively inhibiting the activity of the miRNA on its natural targets and thus suppressing its biological activity[11]. The authors stably overexpressed a miR-10b sponge in breast tumor cells using a viral vector and implanted the cells in mice. This approach gave similar results as the antagomir treatment, demonstrating that the tumor cells, and not the host cells, are the crucial therapeutic targets of the antagomir.

An important limitation of the antagomir therapy reported in this study[2] is that it did not affect the growth of lung metastases once the tumor cells had already disseminated to the lung tissue. Whereas this limits the clinical application of the therapy as dissemination to distant organs is an early event during breast cancer progression[12], it highlights the striking specificity of action of miR-10b in the multi-step process of metastasis. In the future, one may conceive testing miR-10b antagomirs for the prophylactic treatment of tumors that have not yet metastasized, or to counteract progression to metastasis possibly induced by adaptive resistance to antiangiogenic therapy[13].

From the standpoint of developing novel therapies targeting miRNAs, the present study shows that systemic administration of a miR-10b antagomir can effectively target a tumor *in vivo* and affect the biological behavior of a relevant fraction of its cells, is well-tolerated and only accompanied by minor toxicity in mice. Future studies are now warranted to establish its efficacy and safety in more stringent preclinical models, such as spontaneous tumors, and in different types of cancers. Whereas the potential toxicity of targeting any miRNA will be dictated by its specific spectrum of natural targets and expressing tissues, the successful results of systemic miR-10b antagomir treatment supports further development of this strategy to target other miRNAs specifically involved in cancer and metastasis.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

1. Nguyen, D.X., Bos, P.D. & Massague, J. *Nat. Rev. Cancer* **9**, 274–284 (2009).
2. Ma, L. *et al. Nat. Biotechnol.* **28**, X341–347 (2010).
3. Bartel, D.P. *Cell* **136**, 215–233 (2009).
4. Garzon, R., Calin, G.A. & Croce, C.M. *Annu. Rev. Med.* **60**, 167–179 (2009).
5. Nicoloso, M.S., Spizzo, R., Shimizu, M., Rossi, S. & Calin, G.A. *Nat. Rev. Cancer* **9**, 293–302 (2009).
6. Yang, J. *et al. Cell* **117**, 927–939 (2004).
7. Ma, L., Teruya-Feldstein, J. & Weinberg, R.A. *Nature* **449**, 682–688 (2007).
8. Gee, H.E. *et al. Nature* **455**, E8–9; author reply E9 (2008).
9. Baffa, R. *et al. J. Pathol.* **219**, 214–221 (2009).
10. Krutzfeldt, J. *et al. Nature* **438**, 685–689 (2005).
11. Brown, B.D. & Naldini, L. *Nat. Rev. Genet.* **10**, 578–585 (2009).
12. Husemann, Y. *et al. Cancer Cell* **13**, 58–68 (2008).
13. Bergers, G. & Hanahan, D. *Nat. Rev. Cancer* **8**, 592–603 (2008).

Stem cell biologists sure play a mean pinball

Dhruv Sareen & Clive N Svendsen

Mouse fibroblasts are reprogrammed to functional neurons by expression of a few transcription factors.

Developmental biologists have traditionally considered lineage commitment and differentiation to be unidirectional and irreversible processes enforced by stable epigenetic patterns. But recent advances in stem cell biology have shown that, at least under certain conditions, mature adult cells are capable of changing cell type in response to ectopic cues. In the latest example, Vierbuchen *et al.*[1] report in *Nature* that cells from the mesodermal lineage (fibroblasts) can be switched to an ectodermal lineage (neurons) using a simple cocktail of transcription factors. These results promise to open new avenues for the regeneration of neural tissues and provide further evidence that lineage reprogramming might be harnessed to create a variety of desired cell types for regenerative medicine and drug discovery.

The idea that each somatic cell has the potential to regenerate an entire organism arose from pioneering studies of cloning, first in amphibians and then in mammals. The recent discovery that ectopic expression of *Oct4*, *Sox2*, *Klf4* and *c-Myc* can push terminally differentiated cells back in time to a pluripotent state[2,3] has both confirmed this basic idea and changed the way that biologists think about cellular differentiation. In particular, the simplicity of creating induced pluripotent stem (iPS) cells with a small number of transcription factors has led to a reassessment of the plasticity of somatic cells.

As iPS cells can be differentiated to virtually any tissue of the body, they represent a new source of autologous tissue for cell therapy. An alternative approach to the generation of desired cell types is lineage reprogramming, in which one type of mature, differentiated cell is transformed into another. Important insights into this process came from studies showing that exocrine pancreatic acinar cells were reprogrammed to endocrine beta-islet cells using the transcription factors Pdx1, Ngn3 and MafA[4], that expression of Atoh1 in nonsensory cochlear cells generated functional auditory hair cells[5] and that B lymphocytes could be converted to macrophages using C/EBP genes[6]. In neural reprogramming, expression of the neural-specific transcription factors Pax6, Ngn2 and Ascl1 (also Mash1) switched astroglial cells to neurons[7], and expression of Nurr1, Ngn2 and Mash1 generated dopaminergic neurons from restricted neural progenitors[8].

Vierbuchen *et al.*[1] began by compiling a list of 19 important transcription factors known to co-regulate the neurogenic program during development. They infected embryonic and 3-day-old tail-tip fibroblasts isolated from mice expressing green fluorescent protein (GFP) under the control of the neural *Tau* gene with viruses expressing all 19 factors and their permutations. By studying different combinations they found that only five genes—*Ascl1, Brn2/4, Myt1l, Zic1* and *Olig2*—were capable of repatterning fibroblasts to GFP-expressing neurons, termed induced neuronal (iN) cells. The only gene absolutely required was *Ascl1*. However, *Brn2* and *Myt1l* or *Zic1* were needed to increase transformation efficiency, maturation and generation of more complex neuronal phenotypes. Similarly, the efficiency of reprogramming to iPS cells can be improved by increasing the number of exogenous transcription factors.

Vierbuchen *et al.*[1] found that neurotransmitters elicited responses from the iN cells that are characteristic of excitatory glutamatergic neurons and inhibitory GABAergic neurons. However, more restricted dopaminergic, cholinergic and serotonergic projection neurons of the central nervous system were not obtained. Importantly, purified iN cells were capable of firing repetitive action potentials and formed functional synapses. Analysis of the synaptic activity showed a preponderance of excitatory cortical neurons. The majority of the new neural cells were born early and over time developed elaborate fibers and specialized electrophysiology. Nonetheless, it remains to be determined whether iN cells made a complete and stable phenotypic switch or whether a partially intact fibroblast epigenetic pattern survived after incomplete neural reprogramming.

One possible shortcoming of this study is that the authors used only embryonic and neonatal tail tip fibroblasts as the starting material. Although there was little evidence of neuronal production from these tissues, neural crest derivatives that migrate from the neural tube into peripheral tissues during development are known to permeate many skin structures[9]. Indeed, they have been shown to persist until adult development as multipotent skin-derived precursors[10]. To determine whether rare neural crest cells rather than fibroblasts responded to Ascl1 and the other reprogramming factors, it would be useful to perform clonal studies and to evaluate fibroblasts for an assortment of neural crest markers. Interestingly, Sox10, an early neural crest marker, was expressed in tail-tip fibroblasts. Nevertheless, the authors did not rule out the possibility that the iN cells could represent dorsal root ganglion or sympathetic neurons of the peripheral nervous system, evidenced by expression of peripherin in some iN cells. If true, this would further suggest the possibility of a neural crest derivative contaminating the primary fibroblast cultures from which the iN cells were generated.

If iN cells do arise from differentiated fibroblasts, it will be important to work out the underlying mechanisms of reprogramming. Both Vierbuchen *et al.*[1] and the authors of a recent commentary[11] suggest that there may have been a direct transformation from the mesodermal to the ectodermal lineage. Alternatively, it is possible that Ascl1 and the other transcription factors pushed the cells back to a transient primitive (but not pluripotent) state before they differentiated down a different lineage. It seems unlikely that a fully differentiated fibroblast switched fate directly to a fully differentiated neuron. Rather, the fibroblast may have first transitioned back to a neural progenitor state. Vierbuchen *et al.*[1] did not examine the role of cell division in the reprogramming process. Does the fibroblast have to divide in the presence of Ascl1 before acquiring a neural fate? This may allow it to 'reset' crucial expression profiles to a more primitive state on which other exogenous transcription factors would act. Examining neural progenitor markers such as Pax6 and Tbr2 and inhibiting proliferation during reprogramming would help address these issues.

Dhruv Sareen and Clive N. Svendsen are at the Regenerative Medicine Institute at the Cedars-Sinai Medical Center, Los Angeles, California, USA. e-mail: dhruv.sareen@cshs.org or clive.svendsen@cshs.org

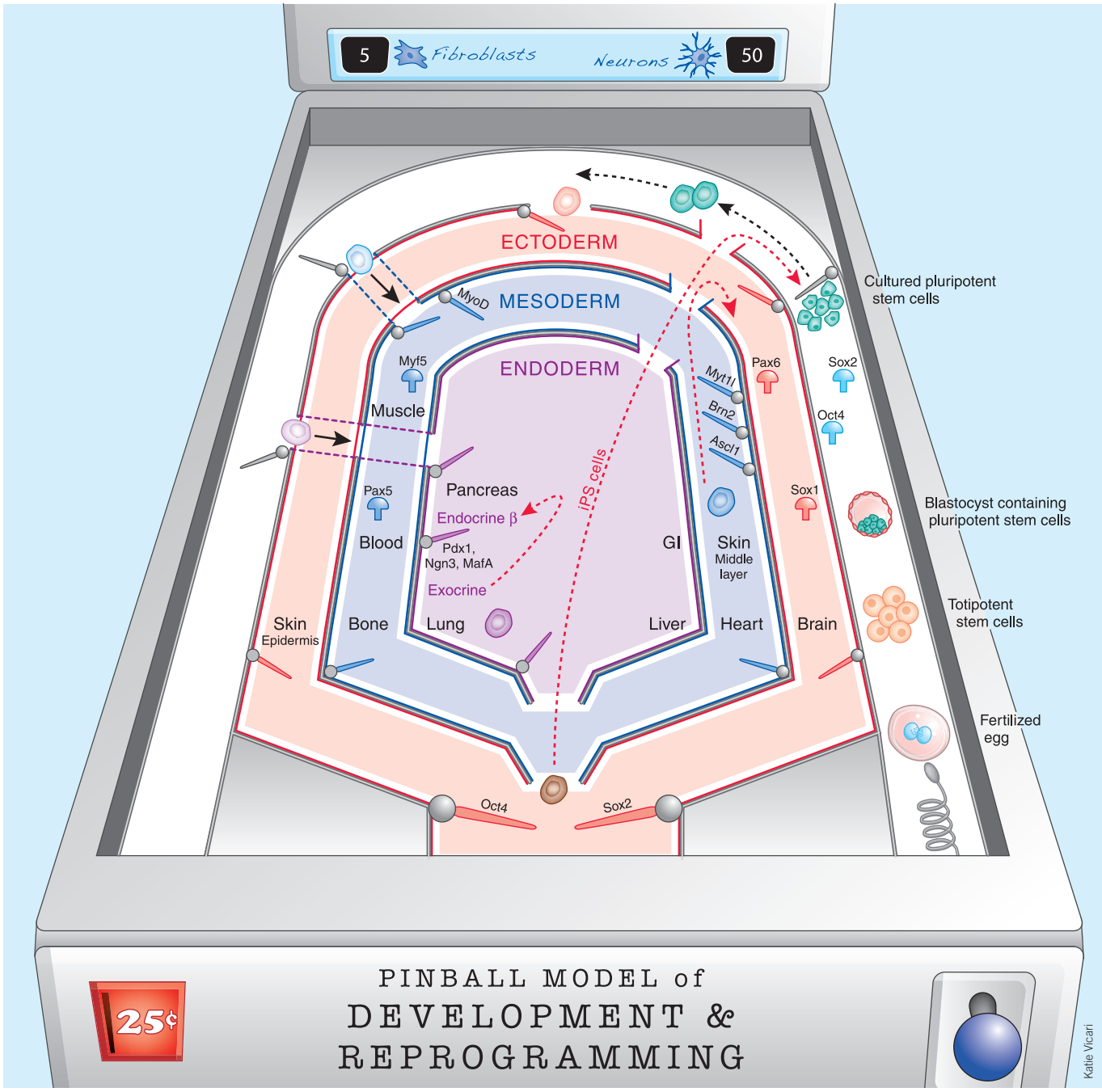


Figure 1 Pinball model of development and reprogramming. The basis of this model is that reprogramming from one cell fate to another requires a dedifferentiation step back to a more primitive state followed by a new path of differentiation. A zygote ('ball') is formed after a sperm enters the egg (bottom right). Totipotent stem cells are produced from the first few divisions of the fertilized egg and become either embryonic or extraembryonic cell types. As development progresses in the 'ball launch lane', embryonic stem cells emerge in the inner cell mass of the blastocyst. The embryonic stem cell expresses endogenous transcription factors ('bumpers') responsible for self-renewal and maintenance of pluripotency. As it moves forward, its developmental potential becomes increasingly restricted. Differentiation into one of the three major dermal lineages is influenced by the developmental guidance cues and epigenetic determinants ('edges') that the cell encounters. Endogenous transcription factors (bumpers) or exogenous transcription factors ('flippers') can drive differentiation forward or, in some cases, flip the cell upward to a less differentiated state (dotted red arrows). A terminally differentiated cell (brown) falling through 'out lanes' can be forced back to pluripotency through 'return lanes' to make iPS cells by overexpression of master regulators (bottom orange flippers) such as Oct4 and Sox2. The expression of other, less powerful transcription factors (smaller flippers) can only tap the cell within a small radius or across an epigenetic edge to another lineage via a transient more primitive state. In an alternative model, a differentiated cell can be pushed directly sideways to another cell type (transdifferentiation) without being flipped to a less differentiated state.

To illustrate the concept of lineage reprogramming via a transient primitive state, we have developed a model based on the metaphor of a pinball machine (**Fig. 1**). In the classic model of C.H. Waddington, a cell moving toward terminal differentiation is represented as a ball rolling down along branching valleys of an epigenetic landscape. In our more interactive model, development of a cell ('ball') begins at the zygote stage and progresses through the blastocyst stage in the 'ball launch lane' (bottom right). As cells move down from the top of the panel, they differentiate into the three germ-cell lineages. A cell is driven to ectoderm, mesoderm or endoderm depending on what cues—'bumpers' or 'flippers'—it strikes. Bumpers represent endogenous transcription factors, whereas flippers represent ectopically expressed transcription factors. The latter can push the cell across epigenetic 'edges' to more primitive stages higher up on the panel or all the way to pluripotency (iPS cells) in the ball launch lane (as shown by *the* pinball wizard, Shinya Yamanaka[2,3]) as well as down specific differentiation pathways. Some master-regulator transcription factors, such as Oct4, can propel the cell a long distance toward pluripotency, whereas lineage-specific transcription factors can only tap it within a single lineage.

In the new work by Vierbuchen *et al.*[1], putative mesodermal fibroblasts were converted to an ectodermal neuronal fate with the Ascl1, Brn2 and Myt1l flippers (blue). For simplicity, we illustrate only two other examples of reprogramming: iPS cells and the intra-lineage conversion of pancreatic endocrine to exocrine cells[4].

In the future it will be of great interest to assess how this model compares to the alternative view that a cell can move laterally on the pinball machine (or through a tunnel in Waddington's model), thus converting directly to a new lineage without passing through a more primitive state. This will be a fundamental principle to understand as the field moves forward.

It will also be important to evaluate the implications of the present study for cell therapy for diseases of the brain. Intra-lineage *in vivo* fate switching within the pancreas or the ear may have immediate impact for diseases like type 1 diabetes or hearing loss. Any newly generated pancreatic beta cells or auditory hair cells would be in a suitable location. However, switching fibroblasts to neurons *in vivo* would be unrewarding as any new neurons produced would not be in their nervous system niche and would therefore be functionally ineffective. On the other hand, *in vitro*–generated neurons or neural progenitor cells may be useful in transplantation-based therapies, and the possibility of generating these cells from patient fibroblasts without going through a pluripotent stage would have the advantage of circumventing tumorigenicity concerns associated with iPS cells. For most clinical applications, however, it would be necessary to be able to expand the cells, as large numbers of cells are typically required for transplantation. Additional work is needed to determine whether the approach of Vierbuchen *et al.*[1] can be modified to produce expandable neural progenitor cells.

Whatever the future holds for this fascinating and fast-moving field, it is apparent that stem cell biologists, in the words of The Who's "Pinball Wizard", sure play a mean pinball.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

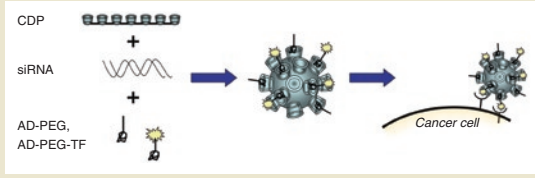
1. Vierbuchen, T. *et al. Nature* **463**, 1035–1041 (2010).
2. Takahashi, K. & Yamanaka, S. *Cell* **126**, 663–676 (2006).
3. Takahashi, K. *et al. Cell* **131**, 861–872 (2007).
4. Zhou, Q., Brown, J., Kanarek, A., Rajagopal, J. & Melton, D.A. *Nature* **455**, 627–632 (2008).
5. Izumikawa, M. *et al. Nat. Med.* **11**, 271–276 (2005).
6. Xie, H., Ye, M., Feng, R. & Graf, T. *Cell* **117**, 663–676 (2004).
7. Berninger, B. *et al. J. Neurosci.* **27**, 8654–8664 (2007).
8. Kim, H.J., Sugimori, M., Nakafuku, M. & Svendsen, C.N. *Exp. Neurol.* **203**, 394–405 (2007).
9. Motohashi, T., Yamanaka, K., Chiba, K., Aoki, H. & Kunisada, T. *Stem Cells* **27**, 888–897 (2009).
10. Toma, J.G., McKenzie, I.A., Bagli, D. & Miller, F.D. *Stem Cells* **23**, 727–737 (2005).
11. Nicholas, C.R. & Kriegstein, A.R. *Nature* **463**, 1031–1032 (2010).

RNA interference in three humans

Many animal studies have shown the therapeutic potential of using small interfering RNAs (siRNAs) to reduce expression of target genes. Although clinical trials with siRNA are underway for a range of diseases[1], it has not yet been demonstrated that delivery of siRNA can trigger RNA interference (RNAi) in humans. For instance, in a clinical trial of intravitreal siRNA for the treatment of blinding choroidal neovascularization, the contribution of non-RNAi mechanisms to the decrease in vascularization could not be eliminated[2].

Now, Davis *et al.*[3] report in *Nature* that siRNA engages the human RNAi machinery to reduce expression of the M2 subunit of ribonucleotide reductase at both the mRNA and protein levels. The study, which is part of a phase 1 clinical trial of systemic siRNA treatment for patients with solid cancers, involved examining biopsies from just three melanoma patients who had received intravenous infusions of siRNA delivered using synthetic nanoparticles.

The nanoparticles (~70-nm diameter) were stabilized by adamantane (AD)-terminated polyethylene glycol (PEG) complexed with a cyclodextrin-based polymer (CDP). Inclusion of the human transferrin (TF) protein on the exposed ends of some of the PEG molecules targeted the nanoparticles to cancer cells expressing the TF receptor.



The authors use 5-nm gold particles to detect the nanoparticles in tumor cells, demonstrating what they believe is the first demonstration of dose-dependent accumulation of systemically delivered targeted nanoparticles in human tumors.

Characterization of the mRNA cleavage products using a modified 5′-RNA-ligand–mediated 'Rapid Amplification of cDNA Ends' method provided mechanistic evidence that the specific siRNAs engaged the RNA interference apparatus. Details of the efficacy of the approach in causing tumor regression have yet to be reported.

Peter Hare

1. Castanotto, D. & Rossi, J.J. *Nature* **457**, 426–433 (2009).
2. Kleinman, M.E. *et al. Nature* **452**, 591–597 (2008).
3. Davis, M.E. *et al. Nature* published online, doi: 10.1038/nature08956 (21 March 2010).

To illustrate the concept of lineage reprogramming via a transient primitive state, we have developed a model based on the metaphor of a pinball machine (**Fig. 1**). In the classic model of C.H. Waddington, a cell moving toward terminal differentiation is represented as a ball rolling down along branching valleys of an epigenetic landscape. In our more interactive model, development of a cell ('ball') begins at the zygote stage and progresses through the blastocyst stage in the 'ball launch lane' (bottom right). As cells move down from the top of the panel, they differentiate into the three germ-cell lineages. A cell is driven to ectoderm, mesoderm or endoderm depending on what cues—'bumpers' or 'flippers'—it strikes. Bumpers represent endogenous transcription factors, whereas flippers represent ectopically expressed transcription factors. The latter can push the cell across epigenetic 'edges' to more primitive stages higher up on the panel or all the way to pluripotency (iPS cells) in the ball launch lane (as shown by *the* pinball wizard, Shinya Yamanaka[2,3]) as well as down specific differentiation pathways. Some master-regulator transcription factors, such as Oct4, can propel the cell a long distance toward pluripotency, whereas lineage-specific transcription factors can only tap it within a single lineage.

In the new work by Vierbuchen *et al.*[1], putative mesodermal fibroblasts were converted to an ectodermal neuronal fate with the Ascl1, Brn2 and Myt1l flippers (blue). For simplicity, we illustrate only two other examples of reprogramming: iPS cells and the intra-lineage conversion of pancreatic endocrine to exocrine cells[4].

In the future it will be of great interest to assess how this model compares to the alternative view that a cell can move laterally on the pinball machine (or through a tunnel in Waddington's model), thus converting directly to a new lineage without passing through a more primitive state. This will be a fundamental principle to understand as the field moves forward.

It will also be important to evaluate the implications of the present study for cell therapy for diseases of the brain. Intra-lineage *in vivo* fate switching within the pancreas or the ear may have immediate impact for diseases like type 1 diabetes or hearing loss. Any newly generated pancreatic beta cells or auditory hair cells would be in a suitable location. However, switching fibroblasts to neurons *in vivo* would be unrewarding as any new neurons produced would not be in their nervous system niche and would therefore be functionally ineffective. On the other hand, *in vitro*–generated neurons or neural progenitor cells may be useful in transplantation-based therapies, and the possibility of generating these cells from patient fibroblasts without going through a pluripotent stage would have the advantage of circumventing tumorigenicity concerns associated with iPS cells. For most clinical applications, however, it would be necessary to be able to expand the cells, as large numbers of cells are typically required for transplantation. Additional work is needed to determine whether the approach of Vierbuchen *et al.*[1] can be modified to produce expandable neural progenitor cells.

Whatever the future holds for this fascinating and fast-moving field, it is apparent that stem cell biologists, in the words of The Who's "Pinball Wizard", sure play a mean pinball.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

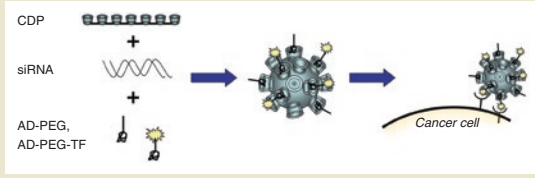
1. Vierbuchen, T. *et al. Nature* **463**, 1035–1041 (2010).
2. Takahashi, K. & Yamanaka, S. *Cell* **126**, 663–676 (2006).
3. Takahashi, K. *et al. Cell* **131**, 861–872 (2007).
4. Zhou, Q., Brown, J., Kanarek, A., Rajagopal, J. & Melton, D.A. *Nature* **455**, 627–632 (2008).
5. Izumikawa, M. *et al. Nat. Med.* **11**, 271–276 (2005).
6. Xie, H., Ye, M., Feng, R. & Graf, T. *Cell* **117**, 663–676 (2004).
7. Berninger, B. *et al. J. Neurosci.* **27**, 8654–8664 (2007).
8. Kim, H.J., Sugimori, M., Nakafuku, M. & Svendsen, C.N. *Exp. Neurol.* **203**, 394–405 (2007).
9. Motohashi, T., Yamanaka, K., Chiba, K., Aoki, H. & Kunisada, T. *Stem Cells* **27**, 888–897 (2009).
10. Toma, J.G., McKenzie, I.A., Bagli, D. & Miller, F.D. *Stem Cells* **23**, 727–737 (2005).
11. Nicholas, C.R. & Kriegstein, A.R. *Nature* **463**, 1031–1032 (2010).

RNA interference in three humans

Many animal studies have shown the therapeutic potential of using small interfering RNAs (siRNAs) to reduce expression of target genes. Although clinical trials with siRNA are underway for a range of diseases[1], it has not yet been demonstrated that delivery of siRNA can trigger RNA interference (RNAi) in humans. For instance, in a clinical trial of intravitreal siRNA for the treatment of blinding choroidal neovascularization, the contribution of non-RNAi mechanisms to the decrease in vascularization could not be eliminated[2].

Now, Davis *et al.*[3] report in *Nature* that siRNA engages the human RNAi machinery to reduce expression of the M2 subunit of ribonucleotide reductase at both the mRNA and protein levels. The study, which is part of a phase 1 clinical trial of systemic siRNA treatment for patients with solid cancers, involved examining biopsies from just three melanoma patients who had received intravenous infusions of siRNA delivered using synthetic nanoparticles.

The nanoparticles (~70-nm diameter) were stabilized by adamantane (AD)-terminated polyethylene glycol (PEG) complexed with a cyclodextrin-based polymer (CDP). Inclusion of the human transferrin (TF) protein on the exposed ends of some of the PEG molecules targeted the nanoparticles to cancer cells expressing the TF receptor.



The authors use 5-nm gold particles to detect the nanoparticles in tumor cells, demonstrating what they believe is the first demonstration of dose-dependent accumulation of systemically delivered targeted nanoparticles in human tumors.

Characterization of the mRNA cleavage products using a modified 5′-RNA-ligand–mediated 'Rapid Amplification of cDNA Ends' method provided mechanistic evidence that the specific siRNAs engaged the RNA interference apparatus. Details of the efficacy of the approach in causing tumor regression have yet to be reported.

Peter Hare

1. Castanotto, D. & Rossi, J.J. *Nature* **457**, 426–433 (2009).
2. Kleinman, M.E. *et al. Nature* **452**, 591–597 (2008).
3. Davis, M.E. *et al. Nature* published online, doi: 10.1038/nature08956 (21 March 2010).

Peptidomimetic antibiotics against pseudomonas

The resistance of many *Pseudomonas aeruginosa* strains to frontline antibiotics complicates the management of hospital-acquired infections. Srinivas *et al*. describe a potential breakthrough in addressing this problem with the identification of a novel mode of bactericidal action. Starting from a cationic antimicrobial peptide, they perform iterative cycles of synthesis and screening of peptidomimetic variants to recover potent antimicrobial compounds. Nanomolar concentrations of two of the optimized compounds, POL7001 and POL7080, are active against the *Pseudomonas* species they test, but not against other bacteria. This selectivity may help to contain the emergence of resistant strains. Instead of lysing bacterial cells, the cyclic 14-amino-acid peptidomimetics appear to bind to and inhibit the outer-membrane protein LptD. This presumably interferes with LptD-mediated incorporation of lipopolysaccharide into the outer leaflet of the cell's outer membrane, a mechanism that does not appear to have been previously targeted in screens for antibiotic activities. Subcutaneous delivery of the compounds within 5 h of bacterial infection is more effective than gentamicin in a mouse model of septicemia. (*Science* **327**, 1010–1013, 2010) *PH*

Piecing together our gut microbiota

A consortium of scientists working under the MetaHIT (Metagenomics of the Human Intestinal Tract) project have sequenced ~577 billion bases of microbial DNA isolated from fecal samples of 124 people from Denmark and Spain. In contrast to previous studies of this kind—which sampled far fewer individuals, sequenced limited regions of DNA (such as the 16S ribosomal DNA) or took a mapping-based approach to match the sequenced fragments of DNA to known bacterial genes—this study used a *de novo* assembly algorithm to piece together billions of short (44 bp or 75 bp) reads into gene-sized stretches of sequence (N50 length of ~2.2 kb). In total, the assembled sequence contained ~3.3 million nonredundant open reading frames, which matched most of the genes of bacteria known to reside in the human gut and identified a core set of shared genes, bacterial species and gene functions present in the sampled individuals. The strategies described in this work for assaying the genetic content of many microbial communities in a deep, unbiased fashion should be useful for investigating questions of both basic research and applied biotechnological value. (*Nature* **464**, 59–65, 2010) *CM*

Tackling HIV's variability

The enormous sequence variety of HIV viruses has stymied vaccine development. A potential solution is the design of mosaic vaccines that maximize the sequence coverage of naturally observed sequence diversity. Barouch *et al*. and Santra *et al*. show that such vaccines can increase the breadth of epitopes that are recognized by T cells in nonhuman primates. Whereas Barouch *et al*. use replication-deficient adenovirus vectors to express mosaic genes for the HIV Gag, Pol and Env genes, Santra *et al*. use mosaic Gag and Nef genes encoded by naked plasmid DNA and vaccinia virus. In both cases, $CD4^+$ and $CD8^+$ T cells recognize a greater range of different epitopes as well as more variations of these epitopes than T cells in animals vaccinated with the consensus or wild-type sequences. Although more experiments will be needed to see whether the increased breadth and depth of the immune response will also be observed in humans and whether it translates into a better protection against viral challenges, the papers present further validation for the mosaic antigen strategy to deal with highly variable pathogens. (*Nat. Med.* **16**, 319–323, 324–328, 2010) *ME*

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

Family ties

The power of familial genomic studies is illustrated in two recent papers in which separate groups sequenced total genomes of multiple members of families with genetic disorders. Using Complete Genomics' (Mountain View, CA, USA) nanoarray sequencing platform, Roach and colleagues sequenced the genomes of two parents and two offspring, both of whom have two genetic disorders, Miller's syndrome and ciliary diskinesia. Looking at both parents and unrelated, affected individuals, the researchers were able to narrow down the critical gene to dihydroorotate dehydrogenase (*DHODH*) in Miller's syndrome and dynein axonemal heavy chain 5 (*DNAH5*) in dyskinesia. Analyzing the pedigrees allowed the researchers to determine that 70 new mutations arose per diploid genome, a mutation rate that differs from previous estimates. Meanwhile, Lupski and colleagues used the SOLID platform (Applied Biosystems; Carlsbad, CA, USA) to sequence ten family members of the first author, who along with three of his siblings, has Charcot-Marie-Tooth disease, a common neuropathy that has been associated with no less than 39 loci. Within those 39 loci, two mutations within a single gene, *SH3TC2* were present in all affected offspring, whereas the parents, as well as some of the unaffected offspring, had only one. Estimating the cost of these genomes is difficult due to fast-moving technological improvements; in the course of one of the studies, the yield increased by a factor of three. (*N. Engl. J. Med.*, published online, doi: 10.1056/nejmoa0908094, 10 March 2010; *Science*, published online, doi: 10.1126/science.1186802, 10 March 2010) *LD*

Giving stroma its due

Stromal cells in the tumor microenvironment modulate the growth characteristics of tumor cells and the efficacy of chemotherapies. Yet the assays and models used in cancer drug discovery do not generally consider tumor-stroma interactions, an omission that may help explain why drugs that appear promising in preclinical studies often fail in clinical trials. Mitsiades and colleagues propose to address this problem with an assay for screening potential cancer drugs in the presence of stromal cells. In their method, tumor cells are labeled with luciferase using retroviral vectors and cultured either with or without unlabeled stromal cells. The cells are then exposed to compounds, and the quantity of viable tumor cells is estimated from the bioluminescence signal. The authors tested a few tumor cell types against several thousand compounds. For more than half the compounds, the presence of stromal cells made the tumor cells less sensitive to the compound, and in a small fraction of cases, sensitivity was increased. These results suggest that including stromal cells in cell-based cancer screens could not only eliminate compounds from the discovery pipeline that would subsequently prove ineffective *in vivo* but also rescue promising compounds that would otherwise be discarded. (*Nat. Med.*, published online, doi:10.1038/nm.2112, 14 March 2010) *KA*

Analyzing 'omics data using hierarchical models

Hongkai Ji & X Shirley Liu

Hierarchical models provide reliable statistical estimates for data sets from high-throughput experiments where measurements vastly outnumber experimental samples.

Interpreting 'omics data often involves statistical analysis of large numbers of loci such as genes, binding sites or single-nucleotide polymorphisms (SNPs). Although the data set as a whole may be rich in information, each individual locus is typically only associated with a limited amount of data. Statistical inference in this context is challenging. A hierarchical model is a useful statistical tool to more efficiently analyze the data, and it is increasingly being used in computational genomics.

A motivating example

Consider a hypothetical microarray experiment with ten genes. For each gene, $\log_2$ expression fold-changes (hereafter referred to as simply 'expression') are observed between tumor and normal tissues in three biological replicates (**Table 1**). To select a gene for follow-up study that is differentially expressed in tumor compared with normal cells, which gene should be the top candidate?

A simple solution is to rank the genes by *t*-statistics

$$t_i = \bar{x}_i / \sqrt{s_i^2 / n}$$

Here n (= 3) is the number of replicates, $\bar{x}_i$ is the average expression of gene i, and s_i^2 is the sample variance. Based on the absolute values of *t*-statistics, gene 2 is the top candidate (**Table 1**).

The data in this example, however, are simulated, with each gene having a 'true' expression μ_i whose measurement is confounded by experimental or biological variability represented by the parameter σ_i^2. (In fact, each expression measurement was randomly drawn from a bell curve–shaped normal distribution with a mean μ_i and variance σ_i^2). The true values of μ_i and σ_i^2, which are unknown to you, are shown in **Table 1**. It turns out the only truly differentially expressed gene is gene 10, which has a nonzero μ_i. Gene 2 thereby represents a false-positive call.

What causes this mistake? Small sample size and the multiplicity of the problem are the reasons. To understand why, it may be helpful to briefly review the key ideas behind statistical inference. The first concept to understand is that of the 'distribution'. Briefly, in the presence of biological or experimental noise and variability, repeated biological measurements are unlikely to be identical, giving rise to a collection, or distribution, of data values. This distribution can be characterized by parameters, such as its mean (or average value) and variance (which quantifies how far the measurements are expected to be from the mean). The parameters are properties associated with infinitely many measurements. In a real scenario, when only a finite number of measurements are available, the true parameter value cannot be obtained. Statistical inference seeks to make statements about the true, also referred to as 'unobserved', parameter value based on the observed data which are called by statisticians as 'samples' drawn from the distribution.

In a *t*-statistic, the sample mean $\bar{x}_i$ represents an estimate of the true mean μ_i of the distribution from which gene i's data are sampled, and the sample variance s_i^2 represents an estimate of the true variance σ_i^2. If the true mean is zero (that is, gene i is not differentially expressed), it is unlikely to obtain a *t*-statistic with a large magnitude.

When the sample size is small, however, the observed sample variance is an unreliable estimate of the true variance of the system. To see why, imagine randomly selecting three data points from a normal distribution with mean 0 and variance 1, which results in the values 0.1, 0.09 and 0.11 (**Fig. 1a**, blue dots). As a result, the observed variance is 0.0001 (or approximately 0) even though the true variance is 1 (that is, much bigger than 0). Another random draw of three data points from the same distribution may give you –1.1, –0.2 and 0.7 (**Fig. 1a**, orange dots) and a totally different observed variance of 0.81. Although the probability that the observed variance significantly deviates from the true variance is small for each individual gene, in a genomic study with many genes, the chance to encounter such deviants for some genes is high.

Small sample variances obtained by chance give rise to large *t*-statistics, which can incorrectly rank nondifferentially expressed genes at the top. This is what happened in our example. The true variance of gene 2 is 1, but the sample variance is 0.005 (**Table 1**); as a result, the *t*-statistic incorrectly ranked gene 2 ($t_2 = 17.5$) on top of the truly differentially expressed gene 10 ($t_{10} = 3.42$). In general, when data analysis involves estimating many parameters or testing many hypotheses but the sample size is small, it is difficult to reliably estimate all parameters or to make correct decisions for all tests simultaneously. This problem is less serious if more samples are available, as more reliable estimates of parameters can be obtained for each gene.

Real gene expression microarray experiments with tens of thousands of genes are examples

Hongkai Ji is in the Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA, and X. Shirley Liu is in the Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Harvard School of Public Health, Boston, Massachusetts, USA.
e-mail: hji@jhsph.edu or xsliu@jimmy.harvard.edu

Table 1 Statistical analysis of example data using either *t*-statistics or a hierarchical model

	Gene, *i*	1	2	3	4	5	6	7	8	9	10
Unobserved parameters	Mean, μ_i	0	0	0	0	0	0	0	0	0	2
	Variance, σ_i^2	2.0	1.0	1.5	0.5	0.7	1.1	1.3	0.9	1.2	1.0
Observed expression data ($\log_2$ fold change)	x_{i1}	0.97	0.73	0.63	1.20	–0.57	3.68	–0.45	1.14	0.34	1.31
	x_{i2}	–0.47	0.78	–0.41	1.48	0.33	–0.68	–0.06	0.40	–0.08	2.59
	x_{i3}	–0.19	0.64	1.93	–0.02	0.26	2.08	–0.74	0.30	1.74	1.03
Gene selection by *t*-statistics	Mean, $\bar{x}_i$	0.10	0.72	0.72	0.89	0.01	1.69	–0.42	0.61	0.67	1.64
	Sample variance s_i^2	0.58	0.005	1.37	0.64	0.25	4.86	0.12	0.21	0.91	0.69
	t_i	0.23	***17.50***	1.06	1.93	0.02	1.33	–2.12	2.32	1.21	3.42
Gene selection by hierarchical model	Adjusted variance, $\hat{o}_i^2$	0.84	0.65	1.09	0.85	0.73	2.22	0.69	0.72	0.95	0.88
	New t_i	0.20	1.53	1.19	1.66	0.01	1.97	–0.87	1.25	1.19	***3.04***

[a]Numerals in bold italics indicate the gene for which the absolute value of the *t*-statistic (t_i) is the largest.

of a 'large *p*, small *n*' problem, where *p* refers to the number of genes and *n* refers to the number of samples. In addition to the multiplicity issue mentioned before, another potential problem is that if the data are not normally distributed, applying a *t*-test can be invalid when the sample size is small[1]. However, this problem is not the focus of the current primer, in which the data in our example are assumed to be normally distributed.

What is a hierarchical model?

One statistical tool for handling large-*p*, small-*n* problems is a hierarchical model. Such a model describes hierarchical relationships between various sources of data variation. The model structure effectively makes it possible to 'borrow' information from all genes to make more reliable statistical inferences about a particular gene. Hierarchical models are conceptually related to regularization techniques, which include Lasso and ridge regression and represent a broad class of methods for handling large-*p*, small-*n* problems (reviewed in refs. 2,3).

In our example, a hierarchical model can be built by assuming that the unobserved mean and variance parameters (that is, μ_i and σ_i^2) of different genes are also sampled from a distribution (denoted as F_0). The distribution is characterized by parameters, such as mean and variance of infinitely many μ_i and σ_i^2 hypothetically collected from different genes. Accordingly, one can imagine that the observed expression data are generated hierarchically by first drawing the mean and variance parameters for each gene from F_0, and then drawing expression measurements for each gene from a gene-specific distribution (that is, a normal distribution with mean μ_i and variance σ_i^2) (**Fig. 1b**).

Naturally, this model describes two sources of variation in the observed expression data. At the top of the hierarchy, the intrinsic similarities and differences between the expression of different genes is mathematically modeled using a distribution (that is, F_0) of the unobserved gene-specific parameters. At the bottom of the hierarchy, the cross-sample variability within a single gene is modeled using a gene-specific distribution with parameters generated from the top-level distribution (**Fig. 1c**). In effect, the top-level distribution describes which gene-specific parameter values are common and which are unusual. The data contain information about the distributions at both levels because there are several replicates for each gene over many different genes.

Although the top-level distribution is usually unknown, it can be estimated using data from the thousands of genes available. Then, using this distribution, the hierarchical model allows one to 'borrow' information across genes to facilitate inference. How much information to borrow is determined by how similar the genes are relative to the cross-sample variability. The intuition is that if the heterogeneity across genes is small, then data from all genes could be informative about the parameters of a particular gene (**Fig. 1c**). Borrowing information across genes essentially increases the effective sample size for making inferences about individual genes[4]. In contrast, the *t*-statistic approach only uses information from a single gene to estimate the mean and variance of the bottom-level distribution for that gene.

Inference using the hierarchical model

The first step in using the hierarchical model is to find a top-level distribution that fits the data (**Fig. 1d**). This process can be intuitively interpreted as learning the cross-gene heterogeneity from the data. The top-level distribution is usually assumed to be a member of a broad family of distributions. In other words, a large number of candidate distributions with the same mathematical form but different parameter values are considered. By varying the parameter values, members in the family are able to describe a variety of distribution patterns of the gene-specific parameters. The analysis starts by finding the distribution (through identifying the parameter value) in the family that fits the data well, and then using the identified distribution to help infer the gene-specific parameters. Commonly used top-level distribution families include 'conjugate priors' and mixtures of simple distributions (e.g., mixture of normal distributions)[5]. The former is typically used if developing a simple computational algorithm is required, and the latter is used if one needs flexibility to describe very complex cross-gene variation patterns.

Next, the top-level distribution is used to adjust the parameter estimate of every gene (**Fig. 1e**). If cross-gene heterogeneity is small, the adjustment will make the parameter estimates of different genes more similar to each other. Here, the hierarchical model borrows from Bayesian inference, a general approach to make statistical inference by combining prior information with observed data[5,6], with the top-level distribution being treated as the prior knowledge about the unobserved mean and variance parameters of individual genes.

Algorithmically, finding the top-level distribution and inferring gene-specific parameters can be implemented simultaneously using standard Bayesian or empirical Bayes techniques[5,6], which sometimes requires advanced and computation-intensive techniques such as Markov chain Monte Carlo[5].

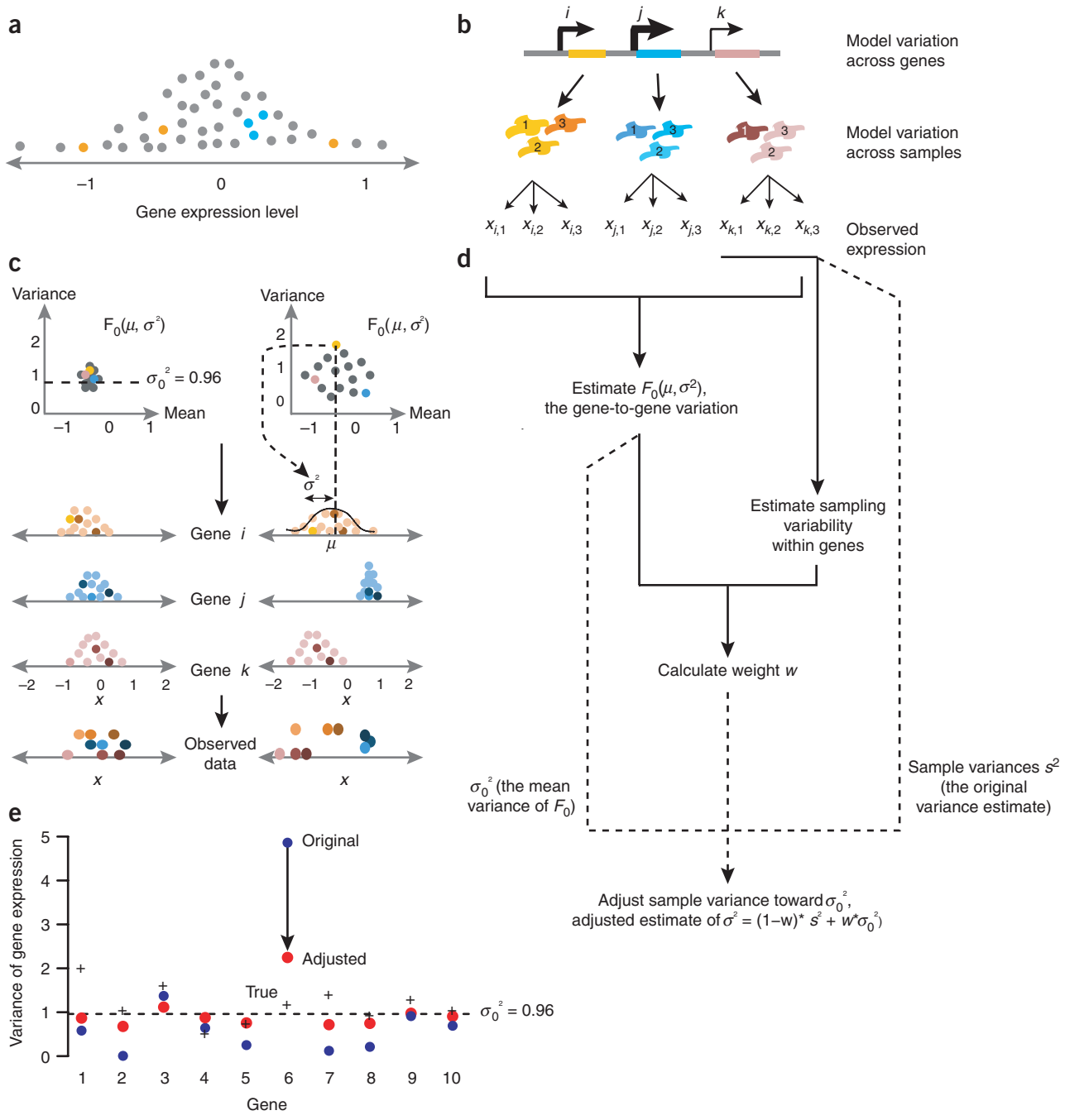


Figure 1 Hierarchical modeling. (**a**) Many analysis techniques, such as *t*-statistics, consider each gene separately. Owing to different sources of biological and experimental variation, if triplicate measurements of the expression of the same gene are collected twice (blue dots and orange dots), the measurements may yield different estimates of the mean and variance of the true distribution that describes the gene's expression (gray). (**b**) A hierarchical model helps produce more reliable estimates of the mean and variance by considering all genes together. It models different sources of biological variation hierarchically. A top-level distribution (F_0) models variation across genes and a bottom-level distribution models variation of the same gene between samples. Data are described by first drawing μ and σ^2 from F_0 for each gene and then drawing expression fold-changes for each gene. (**c**) If different genes have similar mean and variance, data from one gene are informative about the mean and variance of another gene. It is not known a priori whether genes are similar (left, F_0 is tightly clustered) or not (right, F_0 is more spread out). However, this can be learned by looking at the data of many different genes. If genes are similar, the observed gene-to-gene differences can be largely explained by the sampling variability within a gene (bottom, left); otherwise genes are heterogeneous (bottom, right). (**d**) The hierarchical model is applied by first using the observed data to estimate cross-gene variation (that is, F_0), then comparing it to within-gene sampling variability to determine a rule to combine the characteristics shared by all genes with the data specific to each gene for estimating μ and σ^2 (solid lines). In our example, this yields an adjusted variance estimate in the form of a weighted average between the sample variance and the mean of variances σ^2 in F_0 (that is, σ_0^2) (dashed lines). The model was not applied to estimate the gene-specific mean μ. (**e**) The genes' true variances in our example are similar (as in the left side of **c**), which is perceived by the model. As a result, the adjusted variance estimates (red) are closer than the original variance estimates (blue) to the mean variance σ_0^2 (dotted line), which incorporates data from all genes. Overall, the adjusted variance estimates are also closer to the unobserved true variances listed in **Table 1** (black '+').

In our example, applying the hierarchical model yields a new estimate of the variance parameter of a gene. The new estimate of σ_i^2 is a weighted average between the sample variance s_i^2 and the estimated mean variance of all genes (that is, the mean of all variances σ^2 in the estimated F_0, also denoted as σ_0^2) (ref. 7). The sample variance is an estimate of σ_i^2 based on gene *i*'s data, and σ_0^2 represents a shared property of all genes. These two pieces of information are combined using a weight determined automatically by comparing the magnitude of cross-gene variation (with respect to σ_i^2) with that of the within-gene sampling variability (with respect to s_i^2). If the variability among genes is low relative to the sampling variability within a gene, the mean variance σ_0^2 will receive a high weight. On the other hand, if the cross-gene variation is high compared to the within-gene sampling variability, more weight will be given to s_i^2.

The new estimates shift the sample variances toward the common population mean of σ_i^2, and pulls small variances by chance away from zero. Compared with the old estimates s_i^2, the sum of squared error of the new estimates $\hat{\sigma}_i^2$ from the true values is much smaller (3.50 versus 19.46). When the sample variances s_i^2 in the *t*-statistics are replaced by the new estimates, the new *t*-statistics correctly rank gene 10 before gene 2 (**Table 1**). This weighted average technique to estimate the variance is called variance stabilization. It is widely used in analyses of gene expression microarrays[4,8] and chromatin immunoprecipitation on tiling microarrays (ChIP-chip)[7] to detect differentially expressed genes and protein-DNA binding sites, respectively. Naturally, real microarray experiments are more complicated and contain more sources of variation than our example; thus, they can benefit from more sophisticated hierarchical models that capture those types of variation.

The validity of model assumptions, such as those on the hierarchical structure and the distributions at the top and bottom levels, is crucial for the successful application of hierarchical models. When the assumptions hold true, the model brings additional power. Otherwise, the model may not use the information optimally, or may introduce bias that leads to misleading results. Therefore, it is always wise to check the model assumptions by exploring characteristics of the raw data and testing the analysis results using independent information or cross-validation[2].

Other applications

Hierarchical models can be applied to many other problems besides gene expression microarrays and ChIP-chip. For example, in genome-wide association studies, hundreds of thousands of SNPs are tested for association with a phenotype. In a simple scenario, the association can be studied in a linear regression "phenotype = $\alpha_i + \beta_i$ *genotype + noise," where a nonzero coefficient β_i (*i* indexes SNPs) indicates association. With a limited number of samples and many SNPs to evaluate, this approach often lacks the power to distinguish relevant SNPs from random associations. Because SNPs with similar characteristics, such as those that reside in genes in the same pathway or that show a similar degree of evolutionary conservation, have similar potentials to be associated with the phenotype, one can build a hierarchical model to borrow information from similar SNPs to increase the statistical power of association studies[9]. To use this information, one can assume that β_i from different SNPs follow a top-level distribution $N(\mu + \eta^* z_i, \tau^2)$, where z_i is an observed characteristic of SNP *i*, such as conservation score. Here, $\mu + \eta^* z_i$ describes the relationship between a SNP's characteristic and its potential association with the phenotype, and τ^2 describes the heterogeneity among SNPs with the same characteristic. The model can be generalized to incorporate multiple characteristics. One can use data from all SNPs to estimate this top-level distribution (that is, μ, η, τ^2), and make an inference based on new estimates of β_i that combines the top-level distribution with the SNP-specific data.

Application of hierarchical models is not limited to large-*p*-small-*n* data. The models are useful in a broad spectrum of large-*p* problems where the amount of information per locus is limited, with small sample size being a special case. For example, predicting transcription factor binding sites in DNA sequences can be viewed as a problem that probabilistically assigns a 0–1 label to each locus by matching the sequence to a motif model as opposed to a background model. If the sequences are long, there could be random matches to the motif, which leads to false-positive predictions. However, functional transcription factor binding sites tend to cluster in the genome to form regulatory modules. One can build a hierarchical model by assuming that the input sequences consist of background and modules, and the modules in turn consist of background and binding sites, hence binding sites only occur within modules; given the binding site locations, nucleotides are generated according to either the motif or background probability models. Using this hierarchical model, one can first infer the top-level module status by checking sequences from nearby genomic loci, and then combine the module status as prior and the DNA sequence at each locus to infer its binding status. The module status estimated using information across loci helps eliminate many false-positive binding site predictions. In ref. 10, it was shown that the improved estimates of binding site locations increase the power of *de novo* motif discovery.

We conclude by providing two other examples where hierarchical models might be useful yet have not been fully explored. First, if you want to estimate the fold enrichment at ChIP-seq binding loci, but each ChIP and control library has only one replicate sequenced not so deeply, you may estimate a more robust background read count at one locus by borrowing information from other loci. Second, if you want to estimate the binding motif matrices for several transcription factors in the same protein family, but have only a handful of known binding sites for each factor, you can estimate more robust motif matrices by borrowing information across the family. What are other examples? Looking at your own data might reveal the answer.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

1. Ramsey, F.L. & Schafer, D.W. *The Statistical Sleuth: A Course in Methods of Data Analysis* (Duxbury/Thomson Learning; 2002).
2. Hastie, T., Tibshirani, R. & Friedman, J.H. *The Elements of Statistical Learning*, edn. 2 (Springer; 2009).
3. Tibshirani, R. *J. Roy Stat. Soc. B* **58**, 267–288 (1996).
4. Smyth, G.K. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3**, 3 (2004).
5. Gelman, A., Carlin, J.B., Stern, H.S. & Rubin, D.B. *Bayesian Data Analysis* edn. 2 (Chapman & Hall/CRC; 2004).
6. Beaumont, M.A. & Rannala, B. *Nat. Rev. Genet.* **5**, 251–261 (2004).
7. Ji, H. & Wong, W.H. *Bioinformatics* **21**, 3629–3636 (2005).
8. Sartor, M.A. *et al.* BMC *Bioinformatics* **7**, 538 (2006).
9. Chen, G.K. & Witte, J.S. *Am. J. Hum. Genet.* **81**, 397–404 (2007).
10. Zhou, Q. & Wong, W.H. *Proc. Natl. Acad. Sci. USA* **101**, 12114–12119 (2004).

nature biotechnology

Therapeutic silencing of miR-10b inhibits metastasis in a mouse mammary tumor model

Li Ma[1,2], Ferenc Reinhardt[1], Elizabeth Pan[1], Jürgen Soutschek[3], Balkrishen Bhat[3], Eric G Marcusson[3], Julie Teruya-Feldstein[4], George W Bell[1] & Robert A Weinberg[1,2]

MicroRNAs (miRNAs) are increasingly implicated in the regulation of metastasis. Despite their potential as targets for anti-metastatic therapy, miRNAs have only been silenced in normal tissues of rodents and nonhuman primates. Therefore, the development of effective approaches for sequence-specific inhibition of miRNAs in tumors remains a scientific and clinical challenge. Here we show that systemic treatment of tumor-bearing mice with miR-10b antagomirs—a class of chemically modified anti-miRNA oligonucleotide—suppresses breast cancer metastasis. Both *in vitro* and *in vivo*, silencing of miR-10b with antagomirs significantly decreases miR-10b levels and increases the levels of a functionally important miR-10b target, Hoxd10. Administration of miR-10b antagomirs to mice bearing highly metastatic cells does not reduce primary mammary tumor growth but markedly suppresses formation of lung metastases in a sequence-specific manner. The miR-10b antagomir, which is well tolerated by normal animals, appears to be a promising candidate for the development of new anti-metastasis agents.

Ninety percent of cancer-related mortality is caused by metastases formed by disseminated primary tumor cells at distant anatomic sites[1]. Although surgery, radiation therapy and chemotherapy can control many primary tumors effectively, these treatments have limited utility in curbing the metastatic spread of cancer cells and resulting metastasis formation[2]. Critical regulators of the metastatic process, including proteins and miRNAs, are under intensive investigation[2–4]. Understanding the actions of these regulatory molecules provides the basis for molecularly targeted therapeutics. Candidate antimetastasis therapeutic approaches that target tyrosine kinase pathways, the transforming growth factor–β pathway, tumor angiogenesis and the tumor microenvironment have shown efficacy in preclinical studies[5]. Some have been tested in clinical trials: the monoclonal anti-HER2 antibody trastuzumab (Herceptin), when combined with adjuvant chemotherapy, improved metastasis-free survival in women with surgically resected HER2-positive breast cancer[6,7]; bevacizumab (Avastin), a neutralizing antibody against vascular endothelial growth factor, showed measureable but limited benefit in prolonging the time to disease progression in patients with metastatic renal-cell cancer[8] or metastatic colorectal cancer[9]. However, current treatment options rarely cure metastatic cancer. Similarly, there is a lack of prophylactic therapies that are capable of blocking dissemination from primary tumors and preventing future metastasis formation.

Emerging evidence suggests that cancer initiation and progression involve miRNAs—noncoding RNA molecules that act as negative regulators of gene expression. These small cellular RNAs bind to partially complementary sequences at the 3′ untranslated region (UTR) of specific target mRNA molecules, leading to the degradation of the target mRNAs or to the inhibition of their translation, or both[10,11].

Recently, several miRNAs have been found to regulate metastasis[12–17]. For example, we reported that miR-10b is highly expressed in metastatic cancer cells propagated as cell lines as well as in metastatic breast tumors from patients[12]. Its expression is induced by Twist, a transcription factor that orchestrates epithelial-mesenchymal transitions and imparts multiple traits of high-grade malignancy to carcinoma cells[18,19]. miR-10b inhibits translation of the mRNA encoding the homeobox D10 (HOXD10) protein, leading to an increased expression of *RHOC,* a well-characterized pro-metastatic gene[12]. Others reported that ectopic expression of the *BRMS1* (breast cancer metastasis suppressor-1) gene, a negative regulator of Twist expression, leads to decreased expression of miR-10b and RHOC, as well as increased expression of HOXD10 in highly metastatic breast cancer cells[20]. Notably, overexpression of miR-10b in otherwise nonmetastatic breast cancer cells confers invasive and metastatic abilities on these cells when they are growing as xenografts *in vivo*[12].

miR-10b expression levels in unfractionated bulk populations of early-stage tumors removed from breast cancer patients do not predict future metastatic relapse[21]. However, such miRNA expression analyses were carried out on the heterogeneous cell populations present within early primary tumors, in which the invasive and metastatic cells may constitute only a rare subpopulation of the total tumor mass. Activation of Twist and the resulting induction of miR-10b expression may often occur at relatively late stages of primary tumor progression.

miR-10b has indeed been found to be positively associated with high-grade malignancy. This association held true for various cancer types[22]. miR-10b is one of the most significantly upregulated miRNAs in human pancreatic adenocarcinomas[23] and glioblastomas[24], two types of highly metastatic and/or invasive cancers. miR-10b is also

[1]Whitehead Institute for Biomedical Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. [2]MIT Ludwig Center for Molecular Oncology, Cambridge, Massachusetts, USA. [3]Regulus Therapeutics, Carlsbad, California, USA. [4]Department of Pathology, Memorial Sloan-Kettering Cancer Center, New York, New York, USA. Correspondence should be addressed to R.A.W. (weinberg@wi.mit.edu).

Received 11 January; accepted 1 March; published online 28 March 2010; doi:10.1038/nbt.1618

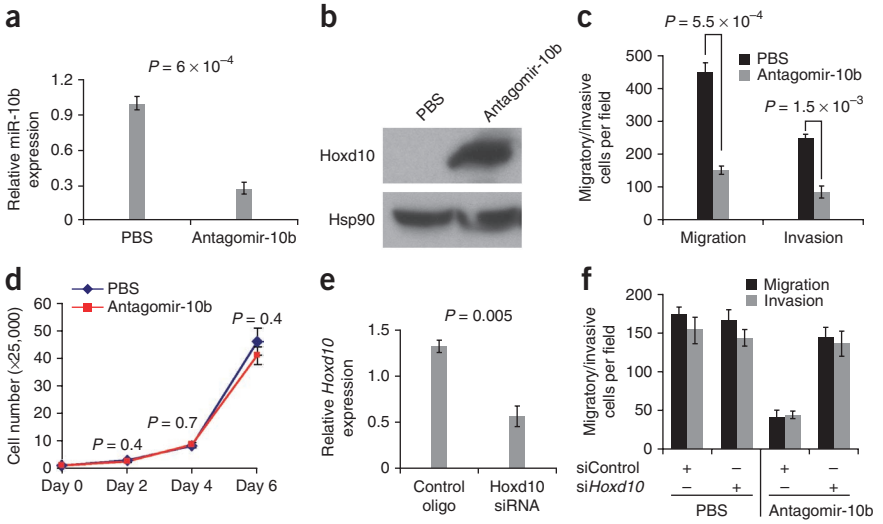


Figure 1 Antagomir-10b can be directly delivered to tumor cells *in vitro* and can inhibit cell motility and invasiveness. (**a**) Real-time RT-PCR of miR-10b in cultured 4T1 cells treated with PBS or antagomir-10b. (**b**) Immunoblotting of Hoxd10 in 4T1 cells treated with PBS or antagomir-10b. Full-length blots and molecular weight markers are presented in **Supplementary Figure 6**. (**c**) Transwell migration assay and Matrigel invasion assay of 4T1 cells treated with PBS or antagomir-10b. (**d**) Growth curves of 4T1 cells treated with PBS or antagomir-10b. (**e**) Real-time RT-PCR of *Hoxd10* in cultured 4T1 cells transfected with *Hoxd10* siRNA or control oligonucleotides. (**f**) Transwell migration assay and Matrigel invasion assay of control siRNA- or *Hoxd10* siRNA-transfected 4T1 cells that are treated with either antagomir-10b or the vehicle (PBS). A representative experiment is shown in triplicate along with s.e.m. in **a**, and **c**–**f**.

upregulated in metastatic hepatocellular carcinomas relative to those that are not metastatic[25]. In human gliomas, miR-10b levels correlate with tumor grade and invasiveness as well as levels of RHOC[26]. This indicates that overexpressed miR-10b may play a causal role in inducing metastatic behavior, but it is unclear whether this miRNA is required for metastasis formation by cancer cells that are naturally highly malignant and, if so, whether it represents a target for the development of antimetastasis therapies—topics that are addressed in the present report.

The development of agents that are directed against miRNAs and are efficacious *in vivo* requires the delivery of these molecules at pharmacologically effective levels. Inhibition of miRNAs can be achieved by antisense oligonucleotides. *In vivo,* the pharmacokinetics and pharmacodynamics of such agents can be improved by chemical modifications designed to enhance their stability and specificity[27]. Several types of antisense-based miRNA inhibitors, including antagomirs, locked nucleic acid oligonucleotides and various types of 2′-*O*-modifed oligonucleotides have proven to be successful for silencing a liver-specific miRNA, miR-122, both in mice[28,29] and in nonhuman primates[30]. However, effective systemic delivery of miRNA antagonists to the neoplastic cells within tumors has not been documented and thus has remained an attractive but untested approach to the development of anticancer agents. Accordingly, we synthesized antagomirs to examine the effects of miR-10b silencing in a 4T1 mouse mammary carcinoma metastasis model[19,31].

RESULTS

Antagomir-mediated silencing of miR-10b in cultured tumor cells

Antagomirs are chemically engineered antisense RNA oligonucleotides against cognate miRNAs[28]. They differ from normal RNAs because of 2′-*O*-methylation of their ribose moieties, partial replacement of phosphodiester bonds by phosphorothioate linkages, and a cholesterol moiety conjugated to the 3′ end[27,28]. Intravenous injection of such antagomirs markedly reduces corresponding miRNA levels in most normal murine tissues except the brain, and the silencing effect can last for over 3 weeks following systemic administration[28]. Whereas the miRNA inhibition by other types of antisense miRNA antagonists reflects either miRNA degradation or competitive inhibition of miRNA detection[32], the decrease in miRNA level after antagomir treatment is likely to reflect miRNA degradation, based on both northern blot analysis under stringent denaturing conditions and successful detection of the degradation products[28].

We first evaluated the ability of the miR-10b antagomir (antagomir-10b) to silence its target miRNA in cultured 4T1 mouse mammary tumor cells, which exhibit high expression levels of both Twist and miR-10b[12,19]. This cell line was isolated as a subpopulation of cells from a mouse mammary tumor with high tumorigenic and metastatic ability, whereas its three isogenic relatives (67NR, 168FARN and 4TO7) are tumorigenic, but are either poorly metastatic or nonmetastatic[19,31].

We treated cultured 4T1 cells with 50 μg/ml antagomir-10b, equivalent to the concentrations used in previous *in vitro* investigations of other antagomirs[33]. Because of the antagomir-induced degradation of its cognate miRNA[27,28], we measured mature miR-10b levels in cellular extracts using a TaqMan RT-qPCR assay, which has been shown to be able to distinguish between similar miRNAs that differ by only a single nucleotide (http://www3.appliedbiosystems.com/cms/groups/mcb_marketing/documents/generaldocuments/cms_042142.pdf). When compared with the vehicle control, cells cultured in the presence of antagomir-10b consistently displayed an ~75% reduction in miR-10b levels (**Fig. 1a**). This coincided with a pronounced induction of the Hoxd10 protein (**Fig. 1b**), whose mRNA is targeted by miR-10b[12]. Hence, antagomir-10b could be readily delivered to cultured cells and could silence miR-10b without use of special transfection procedures.

We performed Transwell migration and Matrigel invasion assays and found that 4T1 cells treated with antagomir-10b displayed a 65–70% decrease in both motility and invasiveness *in vitro* (**Fig. 1c**). In contrast, their *in vitro* proliferation was not affected by this treatment (**Fig. 1d**). In addition, we used a small interfering RNA (siRNA) to knockdown *Hoxd10*, which reduced *Hoxd10* mRNA levels by ~60% (**Fig. 1e**). We then treated 4T1 cells transfected with either control siRNA or *Hoxd10* siRNA with either PBS buffer or antagomir-10b and found that knockdown of *Hoxd10* sufficed to reverse the loss of motility and invasiveness observed in antagomir-10b–treated cells (**Fig. 1f**). Hence, derepressed Hoxd10 expression could explain the reduced cell motility and invasiveness after antagomir-10b treatment.

Pharmacological delivery and specificity of antagomir-10b

Next, we tested the therapeutic efficacy of the miR-10b antagomir when delivered systematically in a tumor metastasis model. To do so, we implanted the 4T1 cells into the orthotopic site—the mammary fat pad—of immunocompetent, syngeneic BALB/c hosts. Consistent

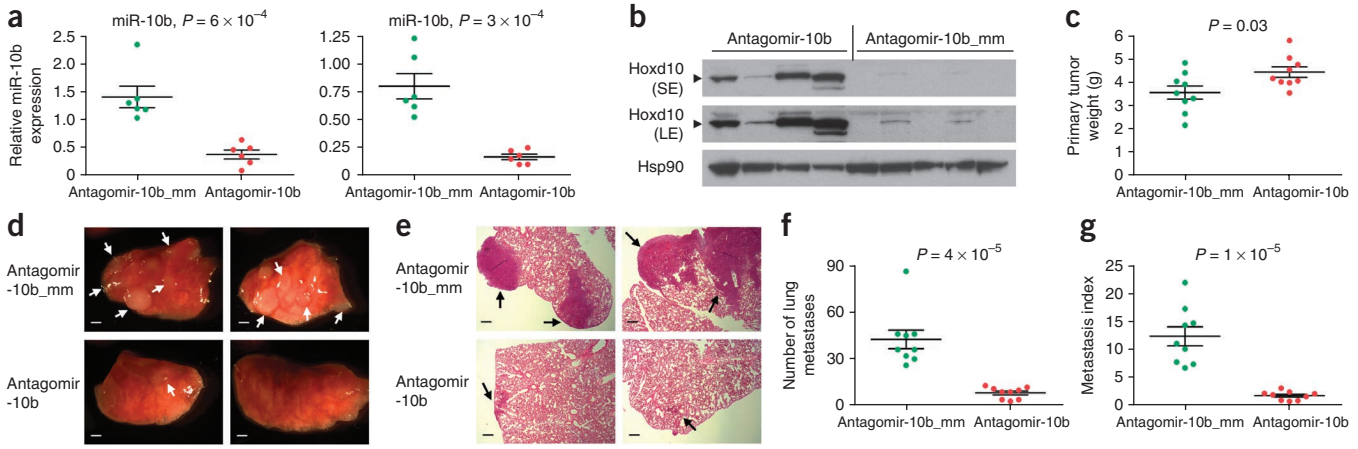


Figure 2 The metastasis-suppressing effect of antagomir-10b is sequence-specific. (**a**) Real-time RT-PCR of miR-10b in primary breast tumors (left panel) and livers (right panel) of 4T1 tumor-bearing mice treated with antagomir-10b or antagomir-10b_mm. Data are presented as mean ± s.e.m. ($n = 6$ mice in each group; each data point represents the mean expression value of triplicates of the sample from one mouse). (**b**) Immunoblotting of Hoxd10 in primary breast tumors of 4T1 tumor-bearing mice treated with antagomir-10b or antagomir-10b_mm. SE: short exposure; LE: long exposure. Full-length blots and molecular weight markers are presented in **Supplementary Figure 6**. (**c**) Primary tumor weight of 4T1 tumor-bearing mice treated with antagomir-10b or antagomir-10b_mm, at 4 weeks after orthotopic implantation. (**d**,**e**) Brightfield imaging (**d**, scale bars, 800 μm) and H&E staining (**e**, scale bars, 200 μm) of the lungs from 4T1 tumor-bearing mice treated with antagomir-10b or antagomir-10b_mm, at 4 weeks after orthotopic implantation. Arrows indicate lung metastases. (**f**,**g**) Number of visible lung metastases (**f**) and metastasis index (= metastasis number divided by primary tumor weight, **g**) in 4T1 tumor-bearing mice treated with antagomir-10b or antagomir-10b_mm, at 4 weeks after orthotopic implantation. Data in **c**, **f** and **g** are presented as mean ± s.e.m. ($n = 9$ mice per group).

with previous reports[19,31], these cells formed primary breast tumors and metastasized to the lungs rapidly. Four weeks after tumor cell implantation, all recipients developed large primary tumors, and multiple visible metastatic nodules could be detected in the lungs with 100% incidence.

In light of the aggressiveness of these 4T1 cells, we began the antagomir treatment 2 d after cancer cell implantation, with the hope of blocking the early steps of the metastatic process, specifically invasion and intravasation. To determine the effective dosage, we referred to the previous *in vivo* testing of multiple antagomir species[27,28,34,35] and applied the following regimen (**Supplementary Fig. 1a**): twice-weekly intravenous doses of 50 mg/kg antagomir for 3 weeks, starting 2 d after tumor cell implantation. At day 28, we euthanized all mice for further analyses.

As gauged by real-time RT-PCR, miR-10b levels were markedly reduced in tissue samples from antagomir-10b–treated mice compared with those from PBS-treated mice: in the liver, which is known to be an organ most accessible to systemically introduced, small RNA-based agents, we observed a 71% reduction ($P = 3 \times 10^{-7}$) of the average miR-10b levels, whereas the average levels of miR-10b were reduced by 65% in the primary breast tumors ($P = 9.9 \times 10^{-5}$) (**Supplementary Fig. 1b**).

To exclude the possibility of nonspecific effects, we synthesized a mutant miR-10b antagomir (termed antagomir-10b_mm) that harbors 12 mismatches within the complementary sequence to miR-10b and does not match any sequence in the mouse genome. Relative to the effect of this mismatch control, the mice injected with antagomir-10b exhibited a 75% reduction of miR-10b levels in primary breast tumors ($P = 6 \times 10^{-4}$) and a 79% reduction of miR-10b levels in their liver ($P = 3 \times 10^{-4}$) (**Fig. 2a**).

Importantly, the levels of the Hoxd10 protein expressed in the primary tumors were markedly increased by antagomir-10b treatment. Unlike antagomir-10b–treated mice, which showed abundant, derepressed Hoxd10 protein expression in primary tumors (**Fig. 2b** and **Supplementary Fig. 1c**), mice treated with antagomir-10b_mm showed low to undetectable Hoxd10 expression in their primary tumors (**Fig. 2b**)—levels comparable to those observed in PBS-treated mice (**Supplementary Fig. 1c**). These results indicated that the miR-10b antagomir, which was readily taken up by the cells in rapidly growing tumors, acts in a sequence-specific manner.

Antagomirs have been shown to be able to discriminate between single-nucleotide mismatches of the targeted miRNA[27]. Accordingly, we examined the levels of other miRNAs in the antagomir-exposed tumors. These analyses indicated that the actions of antagomir-10b are highly specific. Antagomir-10b treatment did not affect the levels of the closely related miR-10a or two unrelated miRNAs, miR-9 and miR-21 (**Supplementary Fig. 2**), both of which are upregulated in clinical breast cancers[36]. Because miR-10a differs from miR-10b by only one nucleotide and is not affected by antagomir-10b treatment, it is highly unlikely that antagomir-10b has a direct effect on any other miRNAs beyond miR-10b. We cannot, however, exclude the possibility that antagomir-10b, by reducing levels of miR-10b, modulates other miRNAs through indirect mechanisms.

Effects on metastasis of systemic antagomir-10b delivery

Reflecting the *in vitro* results (**Fig. 1d**) and our previous observation that miR-10b does not affect tumor cell proliferation[12], we observed no significant difference in primary tumor size between mice treated with PBS or antagomir-10b (**Supplementary Fig. 1d**). In contrast, an 86% decrease in the number of macroscopically visible pulmonary metastases was achieved by antagomir-10b treatment ($P = 5 \times 10^{-5}$). Examination of the lungs revealed an average of 28.6 ± 3.78 visible lesions in mice injected with PBS (**Supplementary Fig. 1e**,**f**), whereas mice treated with antagomir-10b exhibited an average of 4.1 ± 1.6 macroscopic lung metastases (**Supplementary Fig. 1e**,**f**). Presented differently, antagomir-10b–treated mice exhibited an 84% lower ($P = 1 \times 10^{-4}$) metastasis index (metastasis number divided by primary tumor weight, 1.36 ± 0.45) in comparison to the PBS group (8.65 ± 1.25) (**Supplementary**

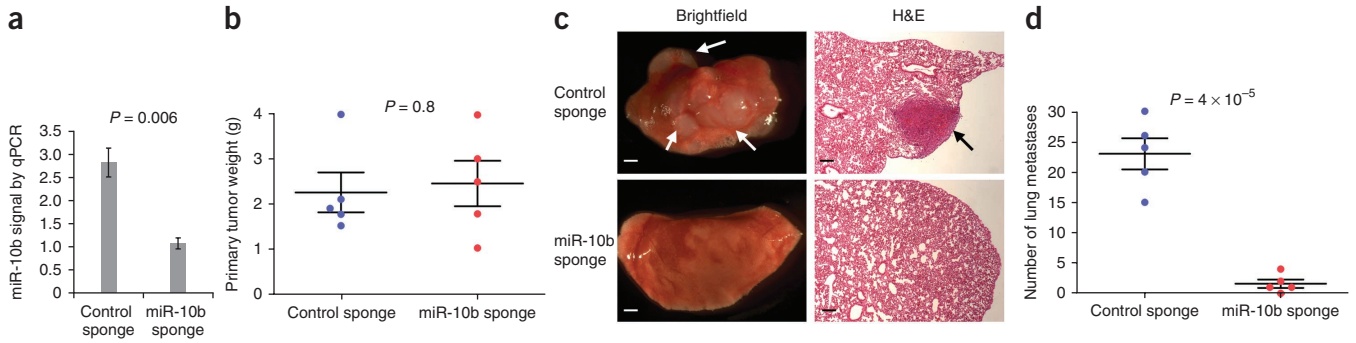


Figure 3 'Sponge'-mediated silencing of miR-10b in tumor cells is sufficient to inhibit metastasis. (**a**) Real-time RT-PCR of miR-10b in 4T1 cells infected with the miR-10b sponge or control sponge. A representative experiment is shown in triplicate along with s.e.m. (**b**) Weight of primary mouse mammary tumors formed by 4T1 cells infected with the miR-10b sponge or control sponge. (**c**,**d**) Brightfield imaging and H&E staining of the lungs (**c**) and number of visible lung metastases (**d**) in mice bearing 4T1 cells infected with the miR-10b sponge or control sponge, at 4 weeks after orthotopic implantation. Arrows indicate lung metastases. Scale bars, 800 μm for brightfield imaging; 200 μm for H&E staining. Data in **b** and **d** are presented as mean ± s.e.m. (n = 5 mice in each group).

Fig. 1g). We anticipate that metastatic cells are the main target of the miR-10b antagomir, because miR-10b is expressed at high levels in these cells, whereas it is present at low levels in normal adult tissues. For example, its expression level in the 4T1 mammary tumor was ~15-fold higher than observed in the normal liver (**Supplementary Fig. 3**).

We also plotted the number of lung metastases versus the miR-10b level expressed in the primary breast tumor in individual recipients (**Supplementary Fig. 1h**). This analysis revealed that levels of this miRNA in primary tumors correlated significantly with pulmonary metastasis numbers in both PBS-treated and antagomir-10b–treated groups of mice (R = 0.94; **Supplementary Fig. 1h**), providing further support for the association of miR-10b with high-grade malignancy.

When using the mismatched antagomir as the control, we observed a moderate 19.7% decrease (P = 0.03) in primary tumor size in the antagomir-10b_mm group relative to the antagomir-10b group (**Fig. 2c**). In contrast, both the number and size of lung metastases were remarkably reduced in the antagomir-10b–treated mice (**Fig. 2d**,**e**). Mice exposed to antagomir-10b or antagomir-10b_mm bore an average of 8.3 and 42.6 lung metastases, respectively (81% reduction, $P = 4 \times 10^{-5}$; **Fig. 2f**); this corresponded to metastasis indices of 1.85 and 12.4, respectively (85% reduction, $P = 1 \times 10^{-5}$; **Fig. 2g**). Taken together, systemic delivery of antagomir-10b has a potent and specific metastasis-suppressing effect on these mouse breast cancer cells without having a notable effect on their ability to grow as primary tumors.

Sponge-mediated miR-10b silencing in tumor cells

To further substantiate the idea that the miR-10b antagomir prevents metastasis largely by targeting the tumor cells, rather than by targeting the host microenvironment, we used an alternative strategy to silence miR-10b in 4T1 cells—a retroviral 'miRNA sponge'[17,37]. This construct encodes a *gfp* mRNA that contains in its 3′UTR multiple tandem binding sites for miR-10b[37]. We observed a 62% reduction in miR-10b levels after sponge expression (**Fig. 3a**), which represented either miRNA degradation or competitive inhibition of detection by sustained expression of the miR-10b sponge (or a combination of both mechanisms). However, even in the latter case, the continued presence of the sponge allows it to absorb miR-10b, thereby interfering with binding of miR-10b to its natural target mRNAs.

This knockdown did not affect the size of the primary mammary tumor formed by sponge-infected cells (**Fig. 3b**). It did, however, dramatically reduce the number of lung metastases (>90% reduction, $P = 4 \times 10^{-5}$; **Fig. 3c**,**d**), suggesting that silencing of miR-10b in primary tumor cells is sufficient to inhibit metastasis.

Effects of miR-10b antagomir on late metastatic stages

To determine whether antagomir-10b has any effect on tumor cells that have already disseminated, we performed tail vein injection of 4T1 cells. This route of transplantation circumvents the initial steps of the invasion-metastasis cascade by introducing cancer cells directly into the lung microvasculature. We then treated the recipient mice with antagomir-10b until they became moribund due to lung metastases (**Fig. 4a**). In this setting, despite expected reduction of miR-10b levels in mouse tissues (66% reduction, $P = 5 \times 10^{-5}$; **Fig. 4b**), both the PBS and antagomir group of mice developed similar numbers of lung metastases (P = 0.7; **Fig. 4c**,**d**), suggesting that antagomir-10b does not affect late stages of the metastatic process, specifically the steps after extravasation of disseminated cells into the foreign tissue parenchyma.

Toxicity assessment of the miR-10b antagomir

To assess potential toxicity of antagomir-10b treatment, we exposed normal mice to the antagomir using the same dosing regimen as described in the metastasis study (**Supplementary Fig. 1a**). As anticipated, intravenous delivery of antagomir-10b reduced miR-10b levels in liver tissues by 72%, whereas antagomir-10b_mm did not modulate miR-10b levels relative to PBS (**Fig. 5a**).

All three groups of mice tolerated the procedure well and exhibited normal behaviors, as determined by activity level and grooming behaviors throughout the study. Body weights as well as lung and heart weights were unaffected by antagomir-10b treatment (**Fig. 5b** and **Supplementary Fig. 4a**,**b**). Histopathological examination of the livers revealed no steatosis, portal or lobular inflammation, necrosis, fibrosis, or biliary change in any of the three groups (**Fig. 5c**). The only notable liver change was increased Kupffer cell macrophages in the antagomir-10b_mm group; however, for unknown reasons this effect was much milder in the antagomir-10b group (**Fig. 5c**). The numbers of white blood cells and lymphocytes in the antagomir-10b group of mice showed a slight decrease and were just below the normal range, when compared to both the PBS group and the mismatched antagomir group (**Fig. 5d**,**e**).

Both compounds, antagomir-10b and antagomir-10b_mm, caused a 8–9% increase in liver and spleen size compared to PBS

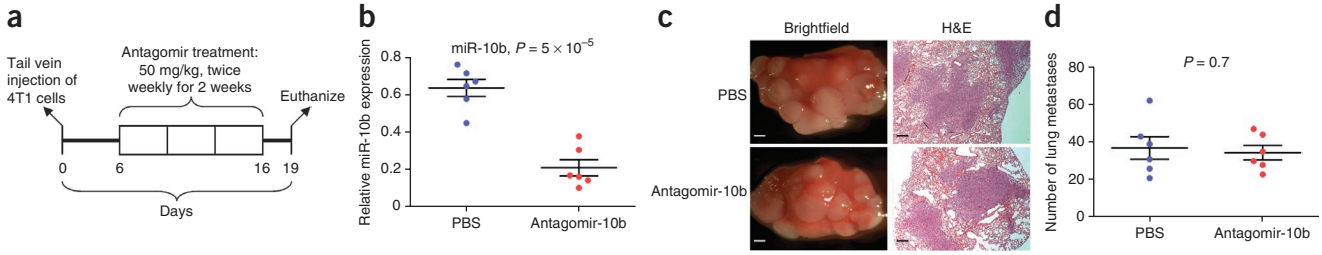


Figure 4 Antagomir-10b treatment does not affect late stages of the metastatic process. (**a**) Schematic representation of the antagomir administration schedule for mice with tail vein injection of 4T1 cells. (**b**) Real-time RT-PCR of miR-10b in livers of mice with tail vein injection of 4T1 cells and subsequent treatment with PBS or antagomir-10b. Data are presented as mean ± s.e.m. ($n = 6$ mice in each group; each data point represents the mean expression value of triplicates of the sample from one mouse). (**c**,**d**) Brightfield imaging and H&E staining of the lungs (**c**) and number of visible lung metastases (**d**) in PBS- or antagomir-10b–treated mice at day 19 after tail vein injection of 4T1 cells. Scale bars, 800 μm for brightfield imaging; 200 μm for H&E staining. Data in **d** are presented as mean ± s.e.m ($n = 6$ mice in each group).

(**Supplementary Fig. 4c**,**d**). Serum chemistry panels revealed unchanged albumin levels (**Supplementary Fig. 5a**). A slight change in several serum proteins and metabolites was observed in both compound groups, but all of them remained in the normal range. Serum levels of the alanine aminotransferase and aspartate aminotransferase liver enzymes were slightly elevated (**Supplementary Fig. 5b**,**c**), and cholesterol and blood urea nitrogen were slightly decreased (**Supplementary Fig. 5d**,**e**). Total bilirubin was the only parameter that showed a marked change, being increased by 2.3-fold upon antagomir-10b treatment (**Supplementary Fig. 5f**), but it nevertheless still remained in the normal range (0–0.9 mg/dl). All these effects were present in both wild-type and mutant antagomir groups (**Supplementary Fig. 5b–f**) and are therefore not specific to miR-10b silencing.

Taken together, when administered at the same dose and frequency as used in the metastasis study, the miR-10b antagomir showed minimal toxic effects in normal animals. The modest effects on liver and spleen size and the levels of several serum proteins and metabolites appear to be related to the chemistry of antagomirs rather than silencing of miR-10b.

DISCUSSION

Our findings demonstrate that targeting of a Twist-induced, metastasis-inducing miRNA is a viable antimetastasis strategy. The metastasis-suppressing effect of antagomir-10b on 4T1 tumor cells *in vivo* phenocopies that of a *Twist* siRNA expressed constitutively in the 4T1 cells[19], suggesting that miR-10b is a functionally important Twist target. Because Twist acts pleiotropically, it is possible that other Twist targets are also essential to the pro-metastatic functions of this transcription factor and that silencing them will also inhibit metastasis.

The actions of antagomir-10b are highly specific, leading to reduction of miR-10b, but not miR-10a or other miRNAs examined. Although systemic administration of the miR-10b antagomir leads to silencing of miR-10b in both metastatic tumor tissues and normal tissues, we reason that metastatic cells are the main target of this agent, because: (i) metastatic tumor tissues express far higher levels of miR-10b than normal tissues; (ii) when administered at the same dose and frequency as used in the metastasis study, the miR-10b antagomir has minimal effects in normal animals; (iii) specific silencing

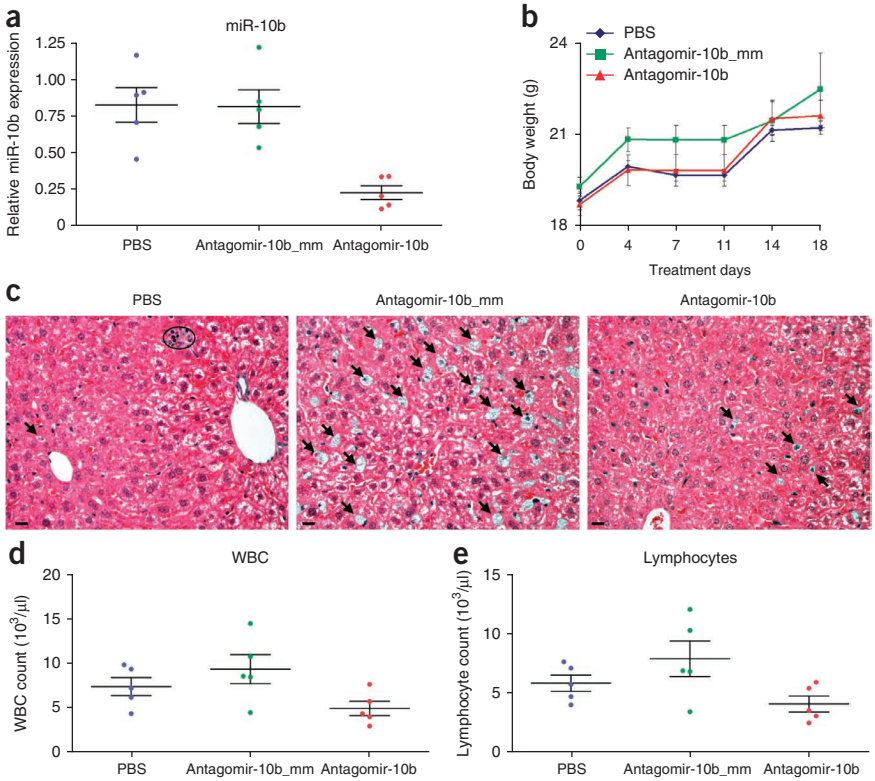


Figure 5 Toxicity assessment following intravenous delivery of antagomir-10b in normal mice. (**a**) Real-time RT-PCR of miR-10b in normal BALB/c mice after treatment with PBS or six doses of 50 mg/kg antagomir-10b or antagomir-10b_mm. Data are presented as mean ± s.e.m. ($n = 6$ mice in each group; each data point represents the mean expression value of triplicates of the sample from one mouse). (**b**) Total body weight was measured twice a week during the study. (**c**) H&E-stained sections of liver samples. Arrows indicate Kupffer cell macrophages. The circle indicates occasional lobular lymphocytes. Scale bars, 30 μm. (**d**,**e**) White blood cell (**d**) and lymphocyte (**e**) count. Data in **b**, **d** and **e** are presented as mean ± s.e.m. ($n = 5$ mice in each group).

of miR-10b in tumor cells is sufficient to phenocopy the antimetastasis effect of systemic antagomir-10b administration.

miR-10a shares the same 'seed' sequence as miR-10b and functions as a pro-metastatic miRNA in pancreatic cancer cells[38]. However, miR-10a expression is not modulated by antagomir-10b treatment and does not compensate functionally for miR-10b loss. The genes of these two miRNAs are located on different chromosomes and their transcription is regulated independently. There is evidence that miR-10a and miR-10b have differential expression patterns, at least in breast cancer cells. For instance, miR-10b is one of the most significantly upregulated miRNAs in the 4T1 metastatic cell line compared to its nonmetastatic or poorly metastatic isogenic relatives (67NR, 168FARN and 4TO7)[39]; in contrast, miR-10a is downregulated in 4T1 and 4TO7 cells compared to 67NR and 168FARN cells[39].

Unmodified antisense oligonucleotides are degraded quickly after systemic administration and have little effect on the miRNA being targeted[28]. This illustrates the need for chemical modifications of oligonucleotides to improve their stability, resistance to RNase and pharmacologic properties. In the present study, the analyses of levels of miR-10b and Hoxd10 suggested that the antagomir effect was still sustained 7 d after the last treatment. This is noteworthy because, unlike normal tissues, these 4T1 tumor cells are actively dividing cells. Because the tissue half-life of antagomirs is 3–3.5 weeks (unpublished observations of J.S., B.B. and E.M.) and because we intravenously administered a relatively high dose of antagomirs repeatedly, it appears that by the time dosing is stopped, the tissue concentration of antagomirs has accumulated to a level well above what is necessary for pharmacological activity. In addition, the antagomir levels achieved in the general circulation could be more than sufficient to titer the targeted miRNA, even in the face of increasing numbers of tumor cells.

Taken together, the actions of the miR-10b antagomir provide a proof-of-principle that antagomirs can be efficiently delivered to rapidly growing tumor cells *in vivo* and can prevent metastasis formation by these highly malignant cells. It remains unclear precisely when cells that are capable of dissemination first arise within primary tumors. Cells with metastatic abilities may already exist at early stages of tumorigenesis[40,41]. Perhaps more commonly, they arise later in the course of multi-step tumor progression[42]. In either case, the clinical utility of inhibiting metastatic dissemination by such cells is limited at present, as dissemination may have already occurred when cancer is detected using current diagnostic methods.

With these reservations in mind, the miR-10b antagomir appears to represent a promising antimetastasis agent that does not act in a cytotoxic fashion on primary tumor cells but instead blocks their ability to launch metastases. We envision that the miR-10b antagomir is a starting point for the development of miRNA-based prophylactic therapies against future metastasis formation. Extensive analyses will be required to determine the long-term efficacy and safety of such agents in various experimental models.

Finally, the functions of miR-10b and the effects of antagomir-10b in other cancer types should be evaluated carefully. Because a single miRNA can potentially target many mRNAs and miRNA-targeted genes themselves may exert differential or even opposing effects in different cellular contexts, the functions of a particular miRNA are often tissue specific, being dependent on the expression pattern of its target mRNAs in a given cell type. In fact, several miRNAs have been reported to be capable of both promoting and suppressing tumorigenesis in a tissue type–dependent fashion[43–46]. For this reason, it remains unclear at present whether agents such as the presently described antagomir will have widespread utility in cancer treatment, or whether such agents will prove to be useful in only a limited set of the tumors encountered in the oncology clinic.

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 C.L. Daige at Regulus for technical assistance, M. Ebert, P. Sharp and S. Valastyan for advice on miRNA sponge design, the Histology Core Labs at Massachusetts Institute of Technology (MIT) and Memorial Sloan-Kettering Cancer Center for assistance with histology, B. Bierie and other members of the Weinberg Lab for useful discussions. L.M. is a recipient of a Life Sciences Research Foundation Fellowship, a Margaret and Herman Sokol Award and a National Institutes of Health (NIH) Pathway to Independence Award (K99/R00). R.A.W. is an American Cancer Society Research Professor and a D.K. Ludwig Cancer Research Professor. This research is supported by an NIH grant to R.A.W. and the Ludwig Center for Molecular Oncology at MIT.

AUTHOR CONTRIBUTIONS

R.A.W. supervised research. L.M., J.S. and E.G.M. designed experiments. J.S., B.B. and E.G.M. provided antagomirs and conditions. L.M., F.R. and E.P. performed most of the experiments. E.G.M. led toxicity assessment at Regulus. J.T.-F. performed pathological analysis. G.W.B. contributed to graphics and statistical analysis. L.M. and R.A.W. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

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. Fidler, I.J. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *Nat. Rev. Cancer* **3**, 453–458 (2003).
2. Steeg, P.S. Tumor metastasis: mechanistic insights and clinical challenges. *Nat. Med.* **12**, 895–904 (2006).
3. Ma, L. & Weinberg, R.A. Micromanagers of malignancy: role of microRNAs in regulating metastasis. *Trends Genet.* **24**, 448–456 (2008).
4. Nicoloso, M.S., Spizzo, R., Shimizu, M., Rossi, S. & Calin, G.A. MicroRNAs—the micro steering wheel of tumour metastases. *Nat. Rev. Cancer* **9**, 293–302 (2009).
5. Steeg, P.S. & Theodorescu, D. Metastasis: a therapeutic target for cancer. *Nat. Clin. Pract. Oncol.* **5**, 206–219 (2008).
6. Romond, E.H. *et al.* Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N. Engl. J. Med.* **353**, 1673–1684 (2005).
7. Piccart-Gebhart, M.J. *et al.* Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N. Engl. J. Med.* **353**, 1659–1672 (2005).
8. Yang, J.C. *et al.* A randomized trial of bevacizumab, an anti-vascular endothelial growth factor antibody, for metastatic renal cancer. *N. Engl. J. Med.* **349**, 427–434 (2003).
9. Hurwitz, H. *et al.* Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *N. Engl. J. Med.* **350**, 2335–2342 (2004).
10. Bartel, D.P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297 (2004).
11. He, L. & Hannon, G.J. MicroRNAs: small RNAs with a big role in gene regulation. *Nat. Rev. Genet.* **5**, 522–531 (2004).
12. Ma, L., Teruya-Feldstein, J. & Weinberg, R.A. Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. *Nature* **449**, 682–688 (2007).
13. Huang, Q. *et al.* The microRNAs miR-373 and miR-520c promote tumour invasion and metastasis. *Nat. Cell Biol.* **10**, 202–210 (2008).
14. Tavazoie, S.F. *et al.* Endogenous human microRNAs that suppress breast cancer metastasis. *Nature* **451**, 147–152 (2008).
15. Asangani, I.A. *et al.* MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pdcd4 and stimulates invasion, intravasation and metastasis in colorectal cancer. *Oncogene* **27**, 2128–2136 (2008).
16. Zhu, S. *et al.* MicroRNA-21 targets tumor suppressor genes in invasion and metastasis. *Cell Res.* **18**, 350–359 (2008).
17. Valastyan, S. *et al.* A pleiotropically acting microRNA, miR-31, inhibits breast cancer metastasis. *Cell* **137**, 1032–1046 (2009).
18. Thiery, J.P. Epithelial-mesenchymal transitions in tumour progression. *Nat. Rev. Cancer* **2**, 442–454 (2002).
19. Yang, J. *et al.* Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* **117**, 927–939 (2004).

20. Edmonds, M.D. *et al.* Breast cancer metastasis suppressor 1 coordinately regulates metastasis-associated microRNA expression. *Int. J. Cancer* **125**, 1778–1785 (2009).
21. Gee, H.E. *et al.* MicroRNA-10b and breast cancer metastasis. *Nature* **455**, E8–E9; author reply, **455**, E9 (2008).
22. Baffa, R. *et al.* MicroRNA expression profiling of human metastatic cancers identifies cancer gene targets. *J. Pathol.* **219**, 214–221 (2009).
23. Bloomston, M. *et al.* MicroRNA expression patterns to differentiate pancreatic adenocarcinoma from normal pancreas and chronic pancreatitis. *J. Am. Med. Assoc.* **297**, 1901–1908 (2007).
24. Ciafre, S.A. *et al.* Extensive modulation of a set of microRNAs in primary glioblastoma. *Biochem. Biophys. Res. Commun.* **334**, 1351–1358 (2005).
25. Hao-Xiang, T. *et al.* MicroRNA-9 reduces cell invasion and E-cadherin secretion in SK-Hep-1 cell. *Med. Oncol.* published online, doi:10.1007/s12032-009-9264-2 (2 July 2009).
26. Sasayama, T., Nishihara, M., Kondoh, T., Hosoda, K. & Kohmura, E. MicroRNA-10b is overexpressed in malignant glioma and associated with tumor invasive factors, uPAR and RhoC. *Int. J. Cancer* **125**, 1407–1413 (2009).
27. Krutzfeldt, J. *et al.* Specificity, duplex degradation and subcellular localization of antagomirs. *Nucleic Acids Res.* **35**, 2885–2892 (2007).
28. Krutzfeldt, J. *et al.* Silencing of microRNAs in vivo with 'antagomirs'. *Nature* **438**, 685–689 (2005).
29. Esau, C. *et al.* miR-122 regulation of lipid metabolism revealed by in vivo antisense targeting. *Cell Metab.* **3**, 87–98 (2006).
30. Elmen, J. *et al.* LNA-mediated microRNA silencing in non-human primates. *Nature* **452**, 896–899 (2008).
31. Aslakson, C.J. & Miller, F.R. Selective events in the metastatic process defined by analysis of the sequential dissemination of subpopulations of a mouse mammary tumor. *Cancer Res.* **52**, 1399–1405 (1992).
32. Davis, S. *et al.* Potent inhibition of microRNA *in vivo* without degradation. *Nucleic Acids Res.* **37**, 70–77 (2009).
33. Li, Q.J. *et al.* miR-181a is an intrinsic modulator of T cell sensitivity and selection. *Cell* **129**, 147–161 (2007).
34. Wurdinger, T. *et al.* miR-296 regulates growth factor receptor overexpression in angiogenic endothelial cells. *Cancer Cell* **14**, 382–393 (2008).
35. Yi, R., Poy, M.N., Stoffel, M. & Fuchs, E. A skin microRNA promotes differentiation by repressing 'stemness'. *Nature* **452**, 225–229 (2008).
36. Iorio, M.V. *et al.* MicroRNA gene expression deregulation in human breast cancer. *Cancer Res.* **65**, 7065–7070 (2005).
37. Ebert, M.S., Neilson, J.R. & Sharp, P.A. MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. *Nat. Methods* **4**, 721–726 (2007).
38. Weiss, F.U. *et al.* Retinoic acid receptor antagonists inhibit miR-10a expression and block metastatic behavior of pancreatic cancer. *Gastroenterology* **137**, 2136–2145 (2009).
39. Dykxhoorn, D.M. *et al.* miR-200 enhances mouse breast cancer cell colonization to form distant metastases. *PLoS ONE* **4**, e7181 (2009).
40. Bernards, R. & Weinberg, R.A. A progression puzzle. *Nature* **418**, 823 (2002).
41. Talmadge, J.E. Clonal selection of metastasis within the life history of a tumor. *Cancer Res.* **67**, 11471–11475 (2007).
42. Scheel, C., Onder, T., Karnoub, A. & Weinberg, R.A. Adaptation versus selection: the origins of metastatic behavior. *Cancer Res.* **67**, 11476–11479; discussion, **67**, 11479–11480 (2007).
43. Krichevsky, A.M. & Gabriely, G. miR-21: a small multi-faceted RNA. *J. Cell. Mol. Med.* **13**, 39–53 (2009).
44. Wang, P. *et al.* microRNA-21 negatively regulates Cdc25A and cell cycle progression in colon cancer cells. *Cancer Res.* **69**, 8157–8165 (2009).
45. Park, J.K., Lee, E.J., Esau, C. & Schmittgen, T.D. Antisense inhibition of microRNA-21 or -221 arrests cell cycle, induces apoptosis, and sensitizes the effects of gemcitabine in pancreatic adenocarcinoma. *Pancreas* **38**, e190–e199 (2009).
46. Huse, J.T. & Holland, E.C. Yin and yang: cancer-implicated miRNAs that have it both ways. *Cell Cycle* **8**, 3611–3612 (2009).

ONLINE METHODS

Antagomir synthesis. Antagomirs were designed and synthesized as described previously[28].

The antagomir-10b sequence was:

5′-oCsoAsoCoAoAoAoUoUoCoGoGoUoUoCoUoAoCoAoGoGsoGsoUsoAs-Chol-3′.

The antagomir-10b_mm sequence was:

5′-oAsoCsoGoAoUoAoAoAoCoGoGoUoUoGoUoCoUoAoCoGsoUsoCsoAs-Chol-3′.

All nucleotides used in synthesis were 2′-*O*-Me-modified (lower case 'o'). Lower case 's' represents a phosphorothioate linkage. 'Chol' represents cholesterol linked through a hydroxyprolinol linkage.

Cell line. The 4T1 cell line was purchased from American Type Culture Collecction and cultured in RPMI medium supplemented with 10% FBS and 1% penicillin/streptomycin. For antagomir treatment, 4T1 cells were incubated with 50 μg/ml antagomir-10b.

RNA interference. SMARTpool Hoxd10 (Dharmacon) represents four pooled SMART-selected siRNA duplexes that target mouse *Hoxd10*. 4T1 Cells were transfected with 200 nM of the *Hoxd10* siRNA or control oligonucleotides using the Oligofectamine reagent (Invitrogen).

miRNA sponge. The miR-10b sponge was constructed using a method modified from previous reports[17,37]: annealed oligonucleotides for tandem miR-10b binding sites were ligated into the pcDNA5-CMV-d2eGFP vector (Invitrogen) digested with XhoI and ApaI. The *gfp* mRNA along with the miR-10b sponge sequence in the 3′UTR was then subcloned into the pBabe-puro vector digested with BamHI and SalI. The control sponge was described previously[17,37].

RNA isolation and miRNA quantification. Total RNA, inclusive of the small RNA fraction, was extracted from cultured cells or homogenized mouse tissues using the mirVana miRNA Isolation Kit (Ambion). Quantification of the mature form of miRNAs was performed using the TaqMan MicroRNA Assay Kit, according to the manufacturer's instructions (Applied Biosystems). The U6 small nuclear RNA was used as an internal control.

Real-time RT-PCR of mRNAs. Total RNA was reverse transcribed with an iScript cDNA synthesis kit (Bio-Rad). The resulting cDNA was used for PCR using the SYBR-Green Master PCR Mix (Applied Biosystems) in triplicates. PCR and data collection were performed on iCycler (Bio-Rad). The expression levels of samples were determined using the standard curve method. Data were normalized to an internal control *Gapdh*. Primer sequences for mouse *Hoxd10* are: forward—AACAGATCTTGTCGAATAGAGCAAC; reverse—GGGCTGTTATTGTACTCTTGGGTTT. Primer sequences for *Gapdh* were described previously[19].

Immunoblotting. Western blot analysis was performed as described previously[12], using antibodies against Hoxd10 (1:200, Santa Cruz) and Hsp90 (1:3000, BD Biosciences).

***In vitro* growth curves.** Growth curves were determined as described previously[12].

***In vitro* migration and invasion assays.** Transwell migration assays and Matrigel invasion assays were performed as described previously[12].

Metastasis studies. Animal experiments were performed in accordance with a protocol approved by the MIT committee on Animal Care. Surgery (mammary fat pad implantation), necropsy and histological analysis were performed essentially as described[12]. Six- to eight-week-old female BALB/c mice (from Jackson Laboratory) were used for implantation of 4T1 cells (one million cells per mouse). The antagomir treatment started 2 d after orthotopic tumor cell implantation: PBS, antagomir-10b, or antagomir-10b_mm (50 mg/kg) was injected via tail vein, twice weekly for 3 weeks. All mice were euthanized 4 weeks after tumor cell implantation due to large primary tumor burdens. The mammary tumors were removed and weighed. Lung metastases were examined and counted under a dissecting microscope equipped with brightfield imaging. Tissue samples were fixed in 10% buffered formalin overnight, washed with PBS, transferred to 70% ethanol and then embedded in paraffin, sectioned and stained with hematoxylin & eosin. To assess the antagomir-10b effect on late stages of the metastatic process, we implanted BALB/c female mice with 5×10^5 4T1 cells via tail vein and started antagomir-10b treatment on day 6 post tumor cell transplantation, with the same dose and frequency as used in the orthotopic experiments; mice were moribund on day 19 owing to lung metastases and were euthanized.

Toxicity assessment. BALB/c mice, five animals per group, were dosed intravenously with PBS or 50 mg/kg of antagomir-10b or antagomir-10b_mm twice a week for 3 weeks (6 doses). Body weight was determined twice a week during the study. Animals were euthanized 24 h after the last dose and tissues were harvested. Blood was collected by retro-orbital bleeding just before euthanasia. An aliquot of whole blood was sent to Molecular Diagnostic Services, Inc., Rabbit & Rodent Diagnostic Associates (RRDA). The remainder was collected in EDTA-treated tubes. Plasma was obtained by removing blood cells through centrifugation and was run on an Olympus Bioanalyzer to determine blood chemistry values. Liver sections were examined by a pathologist (J.T.-F.) for all possible pathological conditions.

Statistical analysis. Data are presented as mean ± s.e.m. Student's *t* test (two-tailed) was used to compare two groups for independent samples, assuming equal variances on all experimental data sets. Spearman's rank correlation test was used to test correlation between a sequence of pairs of values; *R* stands for the correlation coefficient (a measure of the strength of the correlation) that varies from −1 (perfect inverse correlation) through 0 (no correlation) to +1 (perfect positive correlation).

doi:10.1038/nbt.1618

nature biotechnology

V3D enables real-time 3D visualization and quantitative analysis of large-scale biological image data sets

Hanchuan Peng, Zongcai Ruan, Fuhui Long, Julie H Simpson & Eugene W Myers

The V3D system provides three-dimensional (3D) visualization of gigabyte-sized microscopy image stacks in real time on current laptops and desktops. V3D streamlines the online analysis, measurement and proofreading of complicated image patterns by combining ergonomic functions for selecting a location in an image directly in 3D space and for displaying biological measurements, such as from fluorescent probes, using the overlaid surface objects. V3D runs on all major computer platforms and can be enhanced by software plug-ins to address specific biological problems. To demonstrate this extensibility, we built a V3D-based application, V3D-Neuron, to reconstruct complex 3D neuronal structures from high-resolution brain images. V3D-Neuron can precisely digitize the morphology of a single neuron in a fruitfly brain in minutes, with about a 17-fold improvement in reliability and tenfold savings in time compared with other neuron reconstruction tools. Using V3D-Neuron, we demonstrate the feasibility of building a 3D digital atlas of neurite tracts in the fruitfly brain.

Quantitative image analysis is rapidly becoming the bottleneck in high-throughput experiments that seek to gain new insights from advanced 3D microscopic digital images collected using multiple fluorescent probes of different colors, and optionally also for a series of time points[1,2]. This task is often challenging owing to the complexity of the multidimensional image objects in terms of shape (e.g., neurons) and texture (e.g., subcellular organelles), the large scale of the image data, which currently is in the range of hundreds of megabytes to several gigabytes per image stack, the low or limited signal-to-noise ratio of the image data, and the inapplicability of many two-dimensional (2D) image visualization and analysis techniques in these higher dimensional situations.

Real-time 3D visualization-assisted analysis can effectively help an investigator produce and proofread biologically meaningful results in these complicated situations. It is thus highly desirable to have high-performance software tools that can simultaneously visualize large multidimensional image data sets, allow the user to interact and annotate them in a computer-assisted manner, and extract quantitative analysis results.

Many existing software packages, such as Amira (Visage Imaging), ImageJ[3], Chimera[4], Neurolucida (MBF Bioscience) and Image Pro (MediaCybernetics), can be classified into this visualization-assisted analysis category when an image under consideration has a relatively small volume or is primarily processed in a slice-by-slice 2D manner. The Visualization Toolkit[5], in conjunction with the Insight segmentation and registration toolkit[6], is also useful for visualization-assisted analysis of 3D biomedical images. However, the performance of these tools is not yet scalable to multidimensional, multi-gigabyte image data sets.

Here we present a visualization-assisted analysis system called V3D designed to fill this niche. V3D has two distinctive features. First, it contains a visualization engine that is fast enough to render multi-gigabyte, 3D volumetric data in real time on an ordinary computer and to be scalable to very large images. Second, we developed methods to directly pinpoint any XYZ location in an image volume using just one or two mouse clicks while viewing it directly in 3D. The previous alternatives were to do this in a 2D slice or to use a virtual mouse in an expensive stereo-viewing system. These two characteristics in combination make V3D suitable for performing complicated analyses of multidimensional image patterns in a user-efficient way. We demonstrated the strength of V3D by developing a suite of tools for neuroscience, called V3D-Neuron. We applied V3D-Neuron to reconstructing a 3D neuronal structure and produced a 3D digital atlas of stereotypical neurite tracts in a fruitfly's brain.

RESULTS

Real-time 3D visualization of large-scale heterogeneous data

The cross-platform V3D visualization engine renders heterogeneous data including 3D, four-dimensional (4D) and five-dimensional (5D) volumetric image data and a variety of 3D surface objects. These capabilities are demonstrated in **Figure 1a** (and **Supplementary Video 1**), which in a single 3D view shows a multi-channel image stack of a fruitfly brain, a set of surface objects corresponding to brain compartments and several individually reconstructed neurons.

V3D can render image intensity and color in several ways. For the multidimensional intensity data of image stacks, V3D can render a maximum intensity projection, an alpha-blending projection and a cross-sectional view, as well as permitting arbitrary cutting planes through the image stack. For color channels that are often associated with different fluorescent molecules, V3D provides color mapping for improved visualization. V3D supports four categories of 'model' objects: (i) irregular surface meshes (to model objects such as a brain

Janelia Farm Research Campus, Howard Hughes Medical Institute, Ashburn, Virginia, USA. Correspondence should be addressed to H.P. (pengh@janelia.hhmi.org).

Received 30 November 2009; accepted 8 February 2010; published online 14 March 2010; doi:10.1038/nbt.1612

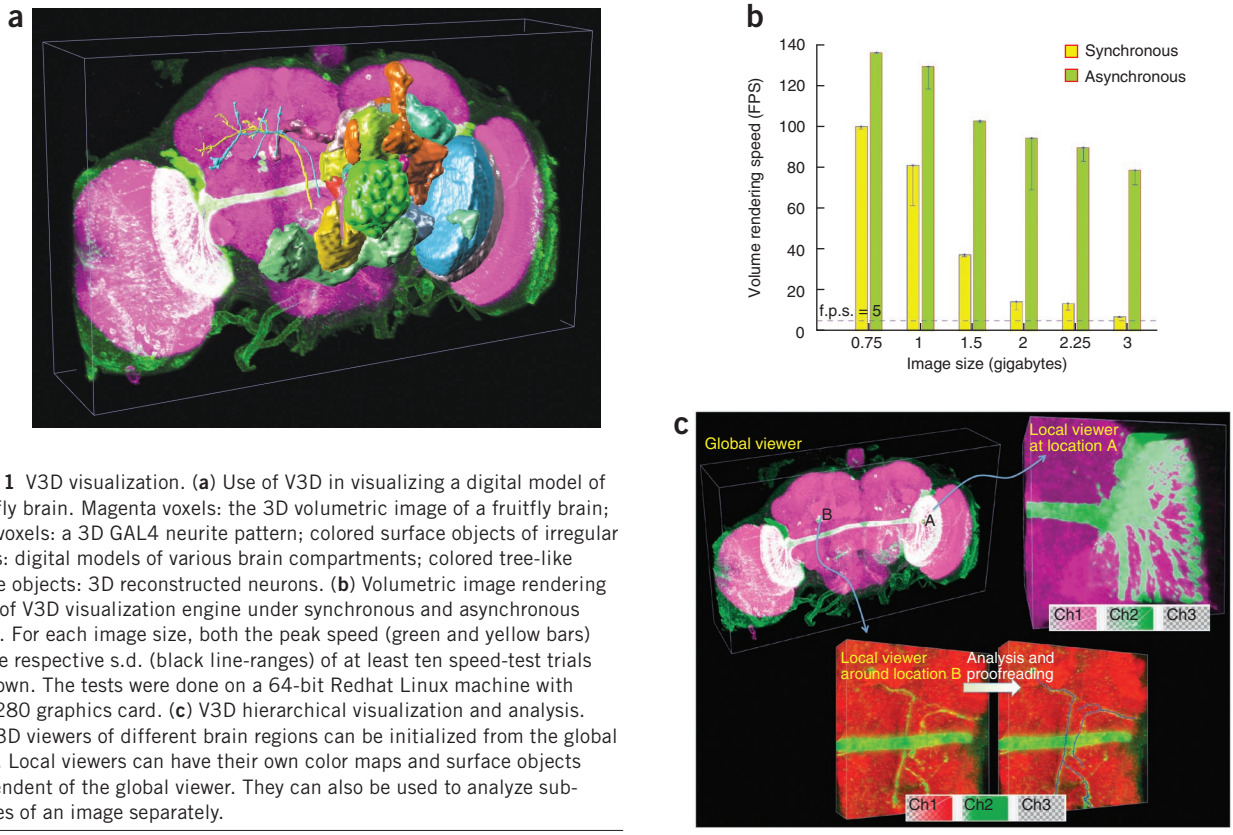


Figure 1 V3D visualization. (**a**) Use of V3D in visualizing a digital model of a fruitfly brain. Magenta voxels: the 3D volumetric image of a fruitfly brain; green voxels: a 3D GAL4 neurite pattern; colored surface objects of irregular shapes: digital models of various brain compartments; colored tree-like surface objects: 3D reconstructed neurons. (**b**) Volumetric image rendering speed of V3D visualization engine under synchronous and asynchronous modes. For each image size, both the peak speed (green and yellow bars) and the respective s.d. (black line-ranges) of at least ten speed-test trials are shown. The tests were done on a 64-bit Redhat Linux machine with a GTX280 graphics card. (**c**) V3D hierarchical visualization and analysis. Local 3D viewers of different brain regions can be initialized from the global viewer. Local viewers can have their own color maps and surface objects independent of the global viewer. They can also be used to analyze sub-volumes of an image separately.

compartment or a nuclear envelope), (ii) spheres (as a simple proxy for globular objects such as cells and nuclei), (iii) fibrous, tubular or network-like structures with complicated branching and connection patterns (to describe the tree-like structure of a neuron, or other relational data or graphs of surface objects), and (iv) markers (to represent various XYZ locations and often used to guide image segmentation, tracing, registration and annotation). Any combination of collections of these objects can be rendered in different colors, styles and in different groupings, on top of an intensity-based multidimensional image. We have used these features to describe a range of digital models of microscopic image patterns, such as a 3D digital nuclear atlas model of *Caenorhabditis elegans*[7] and the fruitfly brain map in **Figure 1a**.

V3D permits interactive rotation, zoom and pan by moving the computer mouse. The speed of the V3D visualization engine is one of the key factors to its usability. It is well known that 3D volumetric image visualization is computationally expensive. To enable real-time 3D rendering of gigabytes of volumetric image data, we first optimized the synchronous rendering of 4D red-green-blue-alpha images using OpenGL 2D or 3D texture mapping (http://www.opengl.org/ and ref. 8), at the full resolution. This resulted in a maximum intensity projection rendering at rates of 15 and 8 frames per second (f.p.s.) in our tests of 2.25- and 3-gigabyte red-green-blue colorimetric images ($1{,}024 \times 1{,}024 \times 768$ and $1{,}024^3$ voxels, respectively) on a 64-bit Linux machine with only 1-gigabyte of graphics card memory (Nvidia GTX280) (**Fig. 1b**). A number of studies show that 5 f.p.s. gives a satisfactory interactive experience[9], indicating that V3D synchronous rendering is sufficiently fast for multi-gigabyte images.

We further designed an asynchronous rendering method performing at almost 80 f.p.s. for a 3-gigabyte image, a tenfold improvement in speed compared to the synchronous rendering (**Fig. 1b**). To do so, we took advantage of the fact that human eyes are not sensitive to the details of a moving object and can distinguish such details only when an object is still[10]. Therefore, in the asynchronous mode, when a user is rotating, or using other ways to interact with a large image, V3D renders a medium-resolution representation of this image. Once such an interaction is over, V3D displays the full resolution. This asynchronous method achieves both high resolution and high speed. It is largely independent of the graphics cards used and is limited by the bandwidth of the peripheral component interconnect express

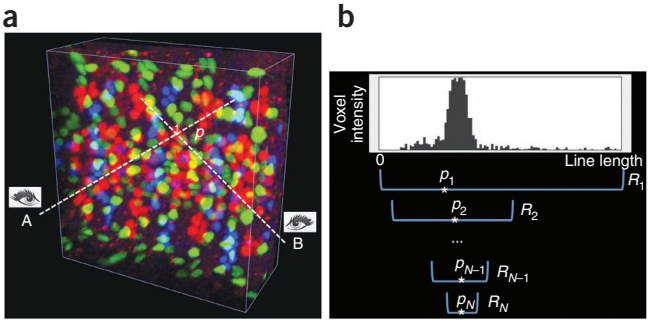


Figure 2 3D pinpointing methods of V3D. (**a**) 3D pinpointing using two mouse-clicks. The color image is a 3D confocal image of neurons, fluorescently tagged for three different transcriptional factors at the same time, in a fruitfly embryo. A and B: non-parallel rays generated at two viewing angles, corresponding to two mouse-clicks; p: the estimated 3D intersection location that is closest to both A and B. (**b**) 3D pinpointing using one mouse-click. p_1 to p_N: the progressively estimated centers of mass; R_1 to R_N: the progressively smaller intervals to estimate p_1 to p_N.

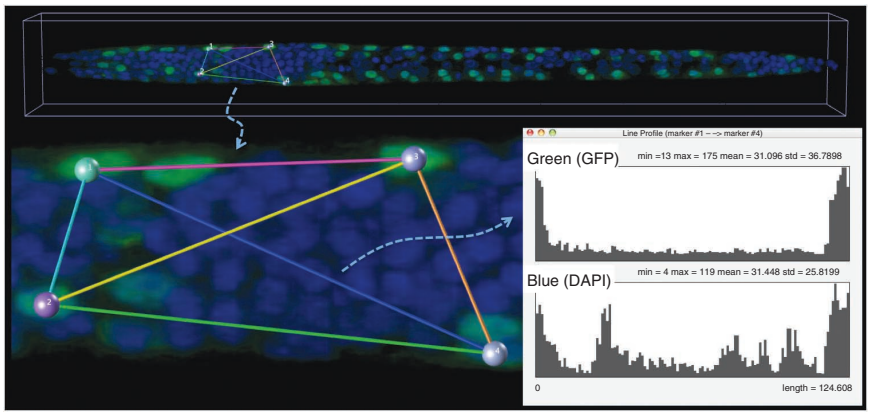


Figure 3 Quantitative measurement of the 3D gene expression level in a *C. elegans* confocal image. Green voxels: myo3: green fluorescent protein (GFP)-tagged body wall muscle cells; blue voxels: DAPI (4,6-diamidino-2-phenylindole)-tagged nuclei for the entire animal; colored spheres: pinpointed markers; colored line-segments: the line-indicator for measuring along different directions and with different starting and ending locations; line profile graph: the channel-by-channel display of the voxel intensity along a line segment.

connection of the underlying computer. We achieved a similar speed improvement on different test machines, such as a Mac Pro with a 512-megabyte graphics card (**Supplementary Fig. 1**).

The only restriction on image size for V3D is that there should be sufficient memory in the underlying hardware. For very large images, for example, 8–16 gigabytes, or for an older machine, the interaction speed becomes an issue. V3D accommodates these situations by displaying two 3D viewers for an image stack: a global viewer and another local viewer of a region of interest dynamically specified by the user (**Fig. 1c** and **Supplementary Video 2a–c**). This combination of dual 3D viewers allows responsive visualization of any detail in a large image stack. Visualization-assisted analysis can be performed in both global and local 3D viewers.

Direct 3D pinpointing of 3D image content

Direct user interaction with the content of a volumetric image requires a user to select one or more points in 3D space, which may be used as 'markers' in subsequent analyses. To complement our fast 3D viewers, we developed two software-based methods to accomplish direct 3D pinpointing within a volumetric image, which has previously only been possible with stereo-viewing hardware or virtual reality equipment.

The first method prompts a user to mouse-click on a point of interest twice while the user is inspecting a volumetric image, each from a different viewing angle (**Fig. 2a** and **Supplementary Video 3a**). Each click defines a ray through the current cursor location orthogonal to the screen. V3D creates a marker at the point in space for which the sum of its Euclidean distance to each ray is minimal (this is to allow for some slight displacement and/or inaccuracy in the user's 2D clicks). This method is independent of the color channels of the image.

The second method pinpoints a 3D marker with just one mouse click (**Fig. 2b** and **Supplementary Video 3b**). Because in this case there is only one ray defined by the click, we estimated the most probable location on the ray that the user intended by looking at the intensity information in a given image color channel. We used the mean-shift algorithm[11] that begins by finding the center of mass (CoM) of the projection ray, and then repeatedly reestimates a CoM using progressively smaller intervals around the proceeding CoM until convergence. This heuristic is very robust in estimating the desired 3D location in real data sets. When there are multiple color channels, the user specifies one by pressing a number key, e.g. 1, 2, while clicking the mouse.

3D pinpointing is useful for measuring and annotating any spatial location in a multi-dimensional image (**Supplementary Fig. 2**), as well as providing manually selected seeds for more complex visualization-assisted analyses and for manually correcting the output of such analyses. For instance, one can quantitatively profile the voxel intensity along the straight line segment connecting a pair of markers. Thus one could quickly measure the gradient of fluorescently labeled gene expression along an arbitrary direction in an image of *C. elegans* (**Fig. 3**). We also used 3D pinpointing to rapidly count the number of neurons in the arcuate nucleus of a mouse brain (**Supplementary Video 4**), and to examine the relative displacement of neurons in a 5D image series of a moving *C. elegans* that was produced by selective plane illumination microscopy (**Supplementary Fig. 3a–c** and **Supplementary Video 5**).

V3D-Neuron: 3D reconstruction of neurites

V3D provides basic software functions for building specialized applications to address specific biological questions. To demonstrate this, we considered digital reconstruction or tracing of 3D neurite structures, which is one of the essential yet bottleneck steps in understanding brain circuit structure and function[12]. We developed V3D-Neuron to trace the 3D structure of a neuron or a neurite tract from images. V3D-Neuron also immediately displays the tracing results superimposed upon the raw image data, letting one proofread and correct the tracing results interactively.

We reconstructed a neuron (**Fig. 4** and **Supplementary Video 6**) based on automated searching of the optimal 'paths' connecting a set of markers, which are locations pinpointed by a user to indicate where the tracing should begin and end (**Fig. 4a**). Our algorithm finds a smooth tree-like structure in the image voxel domain to connect one marker (root) to all remaining markers with the least 'cost', as defined below. We treated individual image voxels as graph vertexes and defined graph edges between the spatial neighboring voxels. The weight of an edge is defined as the product of the inverse of the average intensity of the two neighboring voxels and their Euclidean distance. The least-cost path between a pair of markers will be the one that goes through the highest intensity voxels with the least overall length. We used Dijkstra's algorithm[13] to find the shortest paths (parent-vertex lists) from every non-root marker to the root. We detected in these lists the vertexes where two or more paths merge. These vertexes were treated as the branching points of the reconstructed neuron tree. Subsequently, we represented the traced neuron using the individual segments that connect markers and branching points (**Fig. 4b**), which can be easily edited in 3D whenever needed. For each segment, V3D-Neuron lets an investigator optionally refine its skeleton using a deformable curve algorithm[14] that fits the path as close as possible to the 'midline' of the image region of this segment. In most cases, this not only leads to a smoother path that better approximates the skeleton of a neuron (**Fig. 4c**) but also more accurately estimates the radius of the traced neuron along this skeleton.

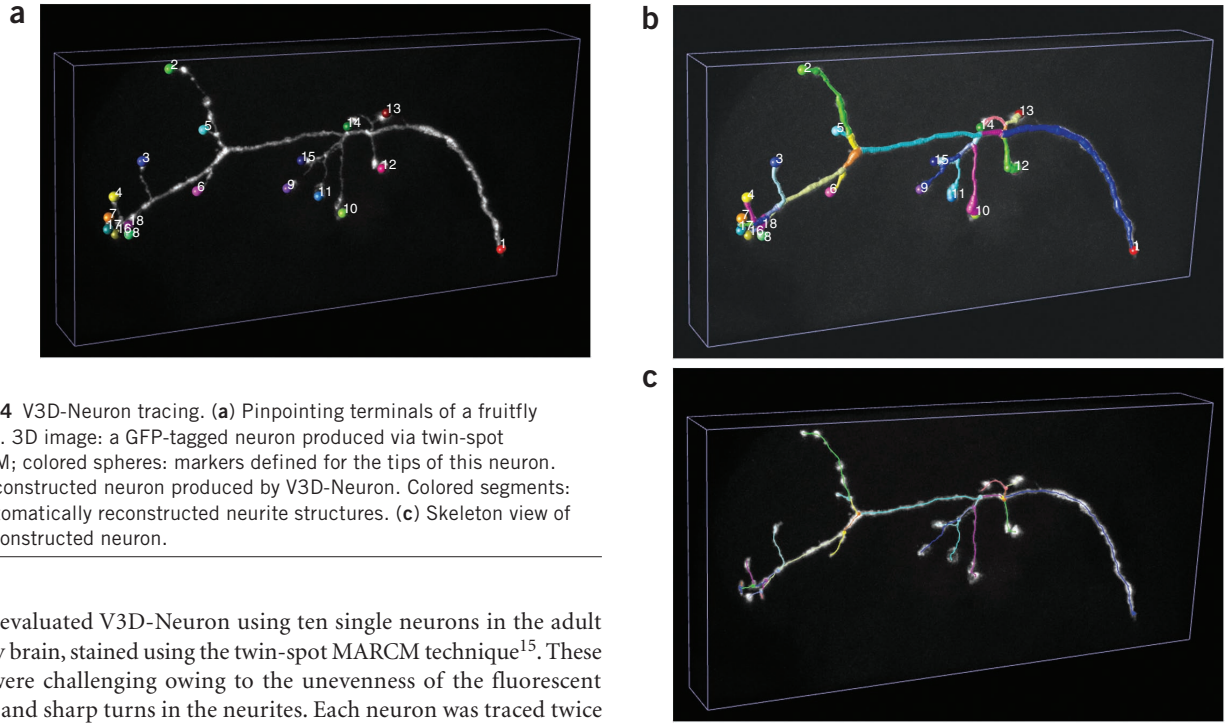


Figure 4 V3D-Neuron tracing. (**a**) Pinpointing terminals of a fruitfly neuron. 3D image: a GFP-tagged neuron produced via twin-spot MARCM; colored spheres: markers defined for the tips of this neuron. (**b**) Reconstructed neuron produced by V3D-Neuron. Colored segments: the automatically reconstructed neurite structures. (**c**) Skeleton view of the reconstructed neuron.

We evaluated V3D-Neuron using ten single neurons in the adult fruitfly brain, stained using the twin-spot MARCM technique[15]. These data were challenging owing to the unevenness of the fluorescent signal and sharp turns in the neurites. Each neuron was traced twice independently. V3D-Neuron allowed us to reconstruct each neuron and correct the potential errors, typically in 3–5 min. Compared to the state-of-the-art 2D manual reconstruction tool of Neurolucida (also with two independent tracings for each neuron), V3D-Neuron avoided the inconsistencies between manual tracings (**Fig. 5a**, e_1–e_4), which were due to human variation in tracing during distinct trials. The difference between V3D-Neuron reconstructions occurred mostly at the tips of the neurons (**Fig. 5a**, e_5) because of the human variation in pinpointing terminal markers. Quantitatively, V3D-Neuron reconstructions have sub-pixel precision, whereas the manual reconstructions have a remarkably bigger variation for the entire structure (**Fig. 5b**). Taking a closer look at the proportion of neuron skeletons that have visible deviation (that is, >2 voxels) between two trials, we found on average only 1.26% of the total path length of a reconstruction was >2 voxels from its counterpart when traced with V3D-Neuron, whereas the same statistic was 21.8% for Neurolucida (**Fig. 5c**). In addition, producing the reconstruction in V3D-Neuron took an order of magnitude less time, as we spent more than 2 weeks obtaining the Neurolucida results. A comparison between V3D-Neuron and other tools (e.g., Image Pro, see **Supplementary Note**) also showed that V3D-Neuron was superior.

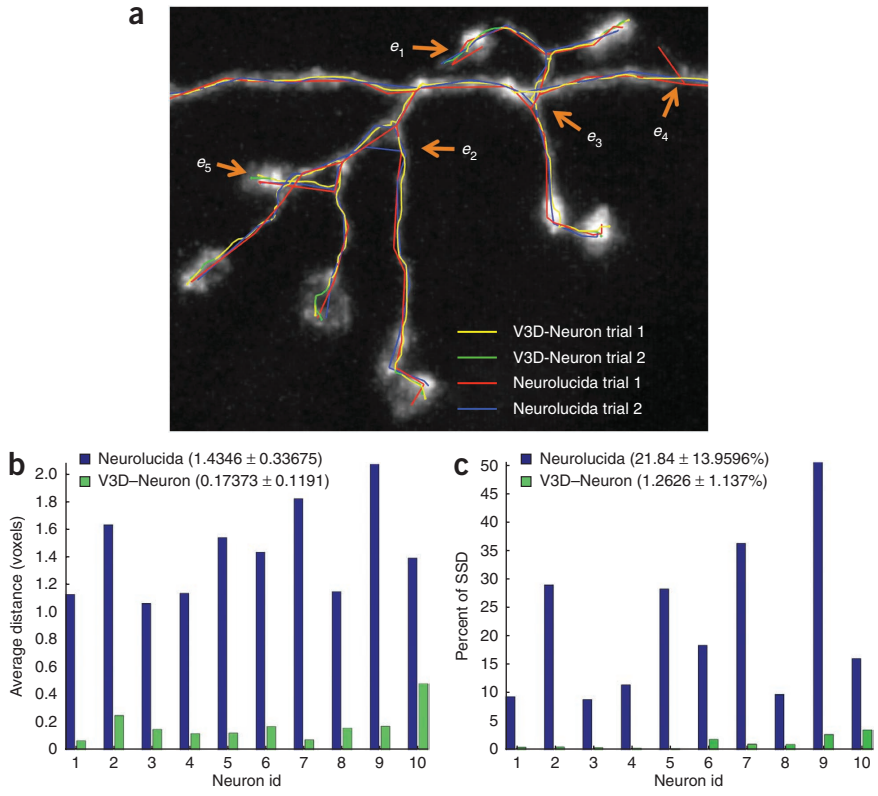


Figure 5 Accuracy of V3D-Neuron reconstructions compared with manual reconstructions. (**a**) Inconsistency of independent trials of reconstructions. e_1, e_2, e_3, e_4: examples of the obvious inconsistent parts in manual reconstructions; e_5: an example of the inconsistent region in V3D-Neuron reconstructions. (**b**) Spatial divergence of reconstructed neurons using different methods, each with two independent runs. Also shown in the legend are the average and the s.d. of the spatial divergence (Online Methods) of all neurons. (**c**) Percent of the neuron structure that noticeably varies in independent reconstructions. Also shown in the legend are the average and the s.d. of this score over all neurons. SSD: substantial spatial distance.

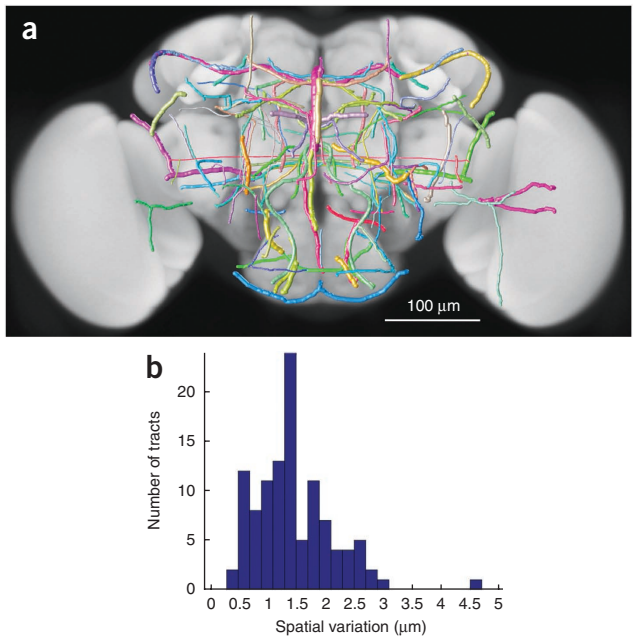


Figure 6 An atlas of stereotyped neurite tracts in a fruitfly brain. (**a**) Statistical models of the 3D reconstructed neurite tracts. Grayscale image: the 'typical' fruitfly brain which was used as the target in 3D alignment; each colored tubular structure: the average of multiple neurite tracts reconstructed from images of the same GAL4 line. The width of each tract equals twice the spatial variation of the respective group of reconstructions. (**b**) Distribution of the spatial variation of all neurite tracts.

A comprehensive 3D atlas of neurite tracts in a brain is critical for understanding both the anatomy and functions of the brain. To demonstrate the feasibility of building such an atlas with V3D-Neuron, we started with more than 3,000 confocal image stacks of the adult fruitfly brain taken of 500 fruitfly enhancer trap GAL4 lines (unpublished data). First, these images were deformably aligned to a 'typical' brain (**Fig. 6a**) using an NC82 intensity pattern in a separate channel (unpublished data), which serves as the reference for accomplishing the automatic alignment. Then, we used V3D-Neuron to trace 111 representative neurite tracts that project throughout the brain (**Fig. 6a** and **Supplementary Video 7**). We reconstructed two to six instances of each tract from multiple replicate images of the same GAL4 line. We produced a mean tract model for each group of reconstructed instances and the average deviation of members of the group from this mean. The overall shape and relative positions of these tracts are reproducible with deviations in the range from 0.5 to 4.5 μm (**Fig. 6b**). This variance approximates the upper bound on the biological variability, as some portion of it was due to alignment optimization differences. Compared to the typical size of an adult fruitfly brain (590 μm × 340 μm × 120 μm), the deviations are small, indicating that the layout of these neural tracts is highly stereotyped in a fruitfly brain.

DISCUSSION

V3D provides an efficient and ergonomic platform for visualization-based bioinformatics in neuroscience, molecular and cell biology, and general biomedical image computing. Compared to many existing visualization-assisted analysis tools, V3D provides real-time visualization of multi-gigabyte volumetric images and the ability to interact with and quantitatively measure the content of such images directly in 3D. Together, these features facilitate rapid computer-assisted analyses and proofreading of the output of such analyses. V3D can also be used for fully automatic analyses of massive image data sets. The software runs on all major computer platforms (Mac, Linux and Windows). It is further enhanced by various built-in V3D toolkits for better visualization and processing, and is extensible by means of user-developed plug-ins (Online Methods). V3D includes a carefully designed graphical user interface and supports common file formats.

The emerging flood of multidimensional cellular and molecular images poses enormous challenges for the image computing community. Although fully automated high-throughput analyses are desired, these approaches are difficult to perfect and are still in their infancy. In this light, it will be helpful to have real-time visualization-assisted tools like V3D to develop, debug, proofread and correct the output of automated algorithms. Indeed, we have already used V3D in developing a number of automatic image analysis methods in addition to V3D-Neuron, including 3D cell segmentation[7], 3D brain registration and 3D neuron structure comparison, among others. Some of these methods, such as the 3D brain registration algorithm, are being deployed in the high-throughput automatic mapping of the 3D neuroanatomy of a fruitfly (unpublished data).

Image analysis is particularly challenging when the signal-to-noise ratio is low or when only part of an image needs to be analyzed, such as reconstructing one neuron from several that are present in an image. The neurite tracing algorithm in V3D-Neuron addresses this challenge by combining global cues supplied by a user and local cues generated automatically from the image data, allowing the user to guide the software when fully automated analysis is difficult. In contrast, existing neuron reconstruction algorithms[16–20] either have not considered such a combination, or are limited to 2D. More importantly, our method is able to complement these algorithms by enabling them to be invoked within V3D, possibly with some guiding marker information, and then their results proofread and corrected in a faster and more accurate way.

Stereotypy, which often refers to the biological invariability, and the degree to which it occurs in the gross and fine-scale anatomy of fly neurons is a question of biological significance and debate[21,22]. The stereotypy of neurite locations found in this work indicates that it may be fruitful to add more neurites to our preliminary atlas and to use these neurites as templates to retrieve and catalog unknown fruitfly neurite patterns. Such information will be useful in modeling 3D anatomical projections and connectivity diagrams of neurons in a brain and in investigating neuron functions. The approach presented here is applicable to a wide range of organisms and at different imaging scales. V3D and V3D-Neuron also have the potential to allow one to manipulate the image data in real-time during functional imaging.

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 work is supported by Howard Hughes Medical Institute. We thank B. Lam, Y. Yu, L. Qu, and Y. Zhuang (Janelia, HHMI) in helping reconstruction of neurites, Y. Yu and L. Qu (Janelia, HHMI) for developing some V3D plug-ins, S. Kim and X. Liu (Stanford) for *C. elegans* confocal images, R. Kerr and B. Rollins (Janelia, HHMI) for the 5D *C. elegans* SPIM images, T. Lee and H. Yu (Janelia, HHMI) for single neuron images, C. Doe (Univ. of Oregon, HHMI) for fruitfly embryo images, A. Jenett (Janelia, HHMI) for fly brain compartments, P. Chung (Janelia, HHMI) for the raw images of fruitfly GAL4 lines, S. Sternson and Y. Aponte (Janelia, HHMI) for the mouse brain image, and K. Eliceiri and C. Rueden (Univ. of Wisconsin,

Madison) for assistance in implementing a V3D plug-in. We also thank G. Rubin, and R. Kerr (Janelia, HHMI) for helpful comments on the manuscript.

AUTHOR CONTRIBUTIONS

H.P. designed this research and developed the algorithms and systems, did the experiments and wrote the manuscript. Z.R. and F.L. helped develop the systems. J.H.S. provided raw images for building the neurite atlas. E.W.M. supported the initial proposal of a fast 3D volumetric image renderer. E.W.M., F.L. and J.H.S. helped write the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

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

1. Peng, H. Bioimage informatics: a new area of engineering biology. *Bioinformatics* **24**, 1827–1836 (2008).
2. Wilt, B.A. *et al.* Advances in light microscopy for neuroscience. *Annu. Rev. Neurosci.* **32**, 435–506 (2009).
3. Abramoff, M.D., Magelhaes, P.J. & Ram, S.J. Image processing with ImageJ. *Biophotonics Int.* **11**, 36–42 (2004).
4. Pettersen, E.F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
5. Schroeder, W., Martin, K. & Lorensen, B. *Visualization Toolkit: An Object-Oriented Approach to 3D Graphics*, 4th edn. (Kitware, Inc, Clifton Park, New York, USA, 2006).
6. Yoo, T.. *Insight into Images: Principles and Practice for Segmentation, Registration, and Image Analysis* (A K Peters, Ltd., Wellesey, Massachusetts, USA, 2004).
7. Long, F., Peng, H., Liu, X., Kim, S. & Myers, E.W. A 3D digital atlas of C. elegans and its application to single-cell analyses. *Nat. Methods* **6**, 667–672 (2009).
8. Wright, R.S. & Lipchak, B.. *OpenGL Superbible*, 3rd edn. (Sams Publishing, Indianapolis, Indiana, 2005).
9. Chen, J.Y.C. & Thropp, J.E. Review of low frame rate effects on human performance. *IEEE Trans. Syst. Man Cybern. A Syst. Hum.* **37**, 1063–1076 (2007).
10. Westheimer, G. Visual acuity. *Annu. Rev. Psychol.* **16**, 359–380 (1965).
11. Fukunaga, K. & Hostetler, L.D. The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Trans. Inf. Theory* **21**, 32–40 (1975).
12. Roysam, B., Shain, W. & Ascoli, G.A. The central role of neuroinformatics in the national academy of engineering's grandest challenge: reverse engineer the brain. *Neuroinformatics* **7**, 1–5 (2009).
13. Dijkstra, E.W. A note on two problems in connexion with graphs. *Numerische Mathematik* **1**, 269–271 (1959).
14. Peng, H., Long, F., Liu, X., Kim, S. & Myers, E. Straightening *C. elegans* images. *Bioinformatics* **24**, 234–242 (2008).
15. Yu, H., Chen, C., Shi, L., Huang, Y. & Lee, T. Twin-spot MARCM to reveal the developmental origin and identity of neurons. *Nat. Neurosci.* **12**, 947–953 (2009).
16. Al-Kofahi, K. *et al.* Rapid automated three-dimensional tracing of neurons from confocal image stacks. *IEEE Trans. Inf. Technol. Biomed.* **6**, 171–187 (2002).
17. Al-Kofahi, K. *et al.* Median-based robust algorithms for tracing neurons from noisy confocal microscope images. *IEEE Trans. Inf. Technol. Biomed.* **7**, 302–317 (2003).
18. Wearne, S.L. *et al.* New techniques for imaging, digitization and analysis of three-dimensional neural morphology on multiple scales. *Neuroscience* **136**, 661–680 (2005).
19. Zhang, Y. *et al.* Automated neurite extraction using dynamic programming for high-throughput screening of neuron-based assays. *Neuroimage* **35**, 1502–1515 (2007).
20. Meijering, E. *et al.* Design and validation of a tool for neurite tracing and analysis in fluorescence microscopy images. *Cytometry* **58A**, 167–176 (2004).
21. Jefferis, G.S. *et al.* Comprehensive maps of *Drosophila* higher olfactory centers: spatially segregated fruit and pheromone representation. *Cell* **128**, 1187–1203 (2007).
22. Murthy, M., Fiete, I. & Laurent, G. Testing odor response stereotypy in the *Drosophila* mushroom body. *Neuron* **59**, 1009–1023 (2008).

ONLINE METHODS

V3D design. We developed V3D using C/C++, with GCC compliers. We used Qt libraries (Nokia) for developing the cross-platform user interface. V3D can be compiled as either a 32-bit or a 64-bit program; the latter can load and save multi-gigabyte and bigger image stacks.

The V3D visualization engine optimizes OpenGL calls to maximize the 3D rendering throughput of a computer graphics card. We rendered volumetric data and surface data separately. For volumetric data, we considered a 4D image stack as the basic display object. A 4D image with any number of color channels is always mapped to red-green-blue-alpha channels for rendering. Then the V3D visualization engine produces either 2D or 3D OpenGL textures, whichever the user chooses, to render the colorimetric 3D volume image. V3D visualization engine also detects automatically a graphics card's data compression capability and lets the user control that. V3D allows a user to switch between synchronous and asynchronous modes for volumetric rendering. In the asynchronous mode, V3D renders a medium-size representation while a user is interacting with the currently displayed volume image, and renders the full-size image once such interaction (e.g., rotating, zooming) is over. The medium-size volume, which has 512 × 512 × 256 red-green-blue voxels by default and can be enlarged if the graphics card has enough memory, is good for visual inspection of image details in most practical cases we have tested. For each image, V3D has a default global 3D viewer and a dynamically generated local 3D viewer for any region of interest. Both the global and local viewers in V3D can use either the synchronous or asynchronous rendering methods for visualization-assisted analysis. Because the local 3D viewer often uses a much smaller volume than the entire image, we often reduced the time of an analysis by aggregating the visualization-assisted analysis results in multiple dynamically invoked local 3D viewers. The volumetric rendering of a 5D image was built upon rendering a series of 4D image stacks, each of which corresponds to a time point. For 3D surface data, the V3D visualization engine uses an object list to manage the rendering of irregular surface meshes (for surface objects of any shape, e.g., brain compartments), spherical objects (for point cloud objects, e.g., cells, nuclei), tubular objects with branches (e.g., neurons, or user-defined line segments) and markers (user-defined XYZ locations). These surface objects can be displayed as solid surface, wiring frame, semi-transparent surface or contours. For the tubular objects with branches, V3D has an additional skeleton display mode. These different visualization methods together make it easy to observe an image and the respective surface objects. They are useful for a number of tasks, for example, developing new image analysis algorithms, proofreading image analysis results, and annotating the image content or surface objects. The speed-test of V3D for **Figure 1b** was done on a Dell Precision T7400 desktop. For the test of synchronous rendering of **Figure 1b**, we used the compression of the graphics card.

A user can develop plug-in programs for V3D. We designed V3D application interfaces (APIs) for writing plug-ins. A plug-in program can request V3D to supply a list of currently opened images, as well as the markers and the regions of interest a user may have defined for these images. After processing any of these data, a plug-in can display the results in either an existing window or a new image viewer window. In this way, it is convenient to use the V3D visualization engine and 3D interaction of image content to develop new image processing tools.

V3D lets one annotate any XYZ location and measure the statistics of image voxels, including the mean intensity, peak intensity, s.d., volume and anisotropy of the shape of the local image patterns in a local region around this 3D location (**Supplementary Fig. 2**). Because for fluorescent images these statistics usually are associated with biological quantities such as gene expression, V3D makes it possible to quickly generate a data table of measurements of gene expression and protein abundance in a number of 3D regions. V3D also supports other types of quantitative measuring, including profiling voxel intensity along a straight line (**Fig. 3**), and profiling the length and the number of branches of a neuron.

V3D-neuron design and algorithm. To model a neuron (or neurite) digitally, we developed a suite of new 3D tracing, 3D visualization, 3D editing and 3D comparison methods on top of V3D. These functions all together are called V3D-Neuron. The version of V3D-Neuron reported here is 1.0.

We modeled a 3D traced neuron as a graph of 'reconstruction nodes (vertexes)'. Each node is described as a sphere centered at a 3D XYZ location p. The diameter of the sphere, r, quantifies the width of the neurite fiber at this location. The value r is defined as the diameter of the largest surrounding sphere centered at p but at the same time within this sphere at least 90% (or 99% at a user's choice) of image voxels are brighter than the average intensity of all image voxels. To describe this graph, we used the SWC format[23]. Through a single run of the tracing algorithm, V3D-Neuron produces a tree graph, or a line graph if there is no branching point. Through multiple sequential reconstruction steps, V3D-Neuron is able to produce a forest of these trees, or looped graphs, or any complicated structures of neurites.

V3D-Neuron tracing method integrates both the shortest path algorithm and the deformable curve algorithm. In V3D-Neuron, we also provided several ways for a user to optionally optimize the performance. We allowed using only the set of image voxels whose intensity is greater than the average voxel intensity of the entire image in the shortest path computation. We also allowed considering only six neighboring voxels along X, Y and Z directions of a $3 \times 3 \times 3$ local cube, instead of using all the 26 neighbors, in formulating the shortest path search. These optimizations substantially accelerate our algorithm. We are also developing other methods, such as detecting the 3D markers automatically, to improve the V3D-Neuron tracing. Although we believe the 3D interactive tracing in V3D-Neuron is a merit but not a caveat, V3D-Neuron of course can also reconstruct a neuron in an off-line (non-interactive) manner, similar to all other conventional off-line automatic neuron tracing methods.

V3D-Neuron visualizes in 3D the reconstructed digital model of a neuron right away. We overlaid a reconstruction on top of the raw image in several ways (**Supplementary Video 8**): (1) display the entire model; (2) display a portion of the model at a user's choice; (3) display only the skeleton of the model; (4) display the model semi-transparently; (5) display the contour of the model. These different methods can also be combined in visualizing and proofreading a reconstruction. In this way, it is straightforward to tell whether or not in the reconstruction there is any error, which can be corrected easily using the 3D editing function below.

In the 3D editing mode (**Supplementary Video 9**), V3D-Neuron automatically represents the entire neuron model as the aggregation of segments, each of which is bounded by a pair of tip or branching nodes. V3D-Neuron renders each segment using a different color, lets a user edit in 3D directly its type (axon, dendrite, apical dendrite, soma or a user-defined type), scale its diameter, and allows deleting it if this segment is deemed to be wrong. Each segment can also be broken into smaller pieces, which can be further edited in the same way. V3D-Neuron can scale, rotate, mirror and translate a neuron in 3D. It also allows annotating a neuron using user-supplied information, and undo/redo operations.

V3D-Neuron can also be used to compare the similarity of multiple neurons. It can display many neurons in the same window in 3D, thus allowing visual comparison of their structures (**Supplementary Video 10**). Moreover, it displays the basic information of the morphology of a neuron, including the total length, the number of branches, number of tips and number of segments. It also provides a method to compute the spatial "distance" (SD) of any two neurons. To compute the distance, we first spatially resampled the neuron model so that the distance between adjacent reconstruction nodes is 1 voxel in the 3D space. Then we computed the directed divergence of neuron A from neuron B, or $D_{DIV}(A,B)$, as the average Euclidean distance of all the resampled nodes in A to B. Finally, the undirected spatial distance between A and B, $SD(A,B)$ is defined as the average of $D_{DIV}(A,B)$ and $D_{DIV}(B,A)$. $SD(A,B)$ is a good indicators of how far away of A and B. However, when A and B are fairly close, $SD(A,B)$ does not well quantify the amount of different portions of the two structures. Therefore, we defined another score called substantial spatial distance (SSD), which is the average distance of all resampled nodes that are apart from the other neuron at least 2 voxels. The percentage of resampled nodes that are substantially distal to the counterpart neuron is a robust indicator of how inconsistent of these two neuron reconstructions. Indeed, for the results in **Figure 5**, the SSD scores of V3D-Neuron reconstructions and manual reconstructions are similar, both about 3 voxels. This is because the major SSD parts of V3D-Neuron reconstructions are at the neuronal terminal regions where the variation of human pinpointing is about 3 voxels, similar to that produced using the Neurolucida tool. However, the percentage of SSD nodes of manual reconstructions is much bigger than that of V3D-Neuron reconstructions (**Fig. 5c**). When two neurons are spatially apart from each

doi:10.1038/nbt.1612

other, SD and SSD scores have comparable values. Of course, when an image is quite dark to see, the 3D pinpointing variation of human subjects becomes more pronounced, which may accordingly enlarge the variation of the paths detected by V3D-Neuron.

Data and software. The multi-platform V3D software can be freely downloaded from http://penglab.janelia.org/proj/v3d. Included in **Supplementary Software** is the 64bit Linux version for Redhat and Fedora. A test image stack and an instruction to run V3D are included in **Supplementary Data**. Additional tutorial movies and test data sets are available at the same web site. The database of stereotyped neurite tracts is provided at http://penglab.janelia.org/proj/flybrainatlas/sdata1_flybrain_neuritetract_model.zip. The entire database can be conveniently visualized using the V3D software itself (**Supplementary Video 10**).

23. Cannon, R.C., Turner, D.A., Pyapali, G.K. & Wheal, W.H. An on-line archive of reconstructed hippocampal neurons. *J. Neurosci. Methods* **84**, 49–54 (1998).

nature biotechnology

Self-sufficient control of urate homeostasis in mice by a synthetic circuit

Christian Kemmer[1], Marc Gitzinger[1], Marie Daoud-El Baba[2], Valentin Djonov[3], Jörg Stelling[1,4] & Martin Fussenegger[1]

Synthetic biology has shown that the metabolic behavior of mammalian cells can be altered by genetic devices such as epigenetic and hysteretic switches[1,2], timers and oscillators[3,4], biocomputers[5], hormone systems[6] and heterologous metabolic shunts[7]. To explore the potential of such devices for therapeutic strategies, we designed a synthetic mammalian circuit to maintain uric acid homeostasis in the bloodstream, disturbance of which is associated with tumor lysis syndrome and gout[8,9]. This synthetic device consists of a modified *Deinococcus radiodurans*-derived protein that senses uric acids levels and triggers dose-dependent derepression of a secretion-engineered *Aspergillus flavus* urate oxidase that eliminates uric acid. In urate oxidase–deficient mice, which develop acute hyperuricemia, the synthetic circuit decreased blood urate concentration to stable sub-pathologic levels in a dose-dependent manner and reduced uric acid crystal deposits in the kidney. Synthetic gene-network devices providing self-sufficient control of pathologic metabolites represent molecular prostheses, which may foster advances in future gene- and cell-based therapies.

State-of-the-art treatment of metabolic disorders consists of small-molecule drug-based interventions to fix out-of-control physiology. However, the daily dosing of drug-based therapies is rather empiric, which may result in undesired side effects. Thus, prosthetic gene-network devices integrated into cells and functionally connected to their metabolism could sense and correct metabolic disturbances as they develop by triggering therapeutic responses in a self-sufficient manner.

As a model system, we explored the potential of a mammalian synthetic circuit for maintaining homeostasis of urate in the bloodstream. Urate is the end-product of purine metabolism, and its homeostasis may be disturbed by genetic predisposition, environmental factors, therapeutic interventions and nutritional imbalances, leading to hyperuricemia. Excess urate favors the formation of pathologic monosodium urate and uric acid crystals in the joints, kidney and subcutaneous tissues, which can trigger a variety of urate-associated pathologies, including the tumor lysis syndrome and gout[9–11].

We began by engineering a synthetic circuit able to sense and respond to uric acid. The circuit uses a bacterial transcriptional repressor (HucR) that binds a DNA sequence motif (hucO) in the absence of uric acid. When uric acid is present, HucR dissociates from DNA, thereby allowing expression of a downstream gene.

HucR and hucO were cloned from *Deinococcus radiodurans* R1 based on molecular insight into this bacteria's remarkable ability to withstand DNA damage caused by ultraviolet radiation and oxidative stress[12]. Recently, a hypothetical uricase regulator (HucR) was identified in *D. radiodurans* and was suggested to play a critical role in the cellular response to oxidative stress[13]. Like mammals, *D. radiodurans* takes advantage of the radical-scavenging activity of uric acid, yet needs to control uric acid levels to prevent crystallization. HucR was shown to bind to a dyad-symmetrical operator site (hucO) in the intergenic region of divergently transcribed *hucR* and a gene encoding a putative uricase, suggesting that both genes are co-repressed by HucR, unless excessive uric acid levels trigger the release of HucR from hucO and induce uricase expression[13,14]. Taken together, these prior observations suggested that HucR and hucO could be used as building blocks in a uric acid–responsive expression network.

The uric acid–responsive expression network (UREX) required engineering HucR and hucO for optimal performance in mammalian cells (**Supplementary Table 1**). The HucR start codon was modified (GTG→ATG) and fused to a Kozak consensus sequence for maximum expression of the repressor protein in mammalian cells (mHucR). mHucR was engineered to be a stronger repressor by fusing it to the C terminus of the Krueppel-associated box (KRAB) protein domain[15]. Notably, fusing mHucR to the N terminus abolished the urate-responsiveness of mHucR (data not shown). The resulting repressor is a chimeric mammalian urate-dependent transsilencer (mUTS). To engineer a promoter that strongly drives expression of a target gene in the presence of uric acid, yet has low basal expression levels in the absence of uric acid, we placed eight tandem hucO modules downstream of a simian virus 40 promoter (P_{SV40}-$hucO_8$, referred hereafter as P_{UREX8}) (**Fig. 1a**). P_{UREX8} was chosen as part of the uric acid sensor circuit because it provided the tightest repression in preliminary tests (for a detailed and quantitative evaluation of alternative promoter configurations, see below).

[1]Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland. [2]Institut Universitaire de Technologie, IUTA, Département Génie Biologique, Villeurbanne Cedex, France. [3]Département de Médecine, Université de Fribourg, Fribourg, Switzerland. [4]Swiss Institute of Bioinformatics, ETH Zurich, Zurich, Switzerland. Correspondence should be addressed to M.F. (fussenegger@bsse.ethz.ch) or J.S. (joerg.stelling@bsse.ethz.ch).

Received 1 February; accepted 19 February; published online 28 March 2010; doi:10.1038/nbt.1617

We characterized the performance of P_{UREX8} by cloning it upstream of a reporter gene encoding SEAP (human placental secreted alkaline phosphatase), which is assayed in the culture supernatant. In the absence of uric acid, mUTS binds $hucO_8$ and silences transcription from P_{UREX8} (**Fig. 1b**). However, in the presence of uric acid, mUTS is released from $hucO_8$ thereby inducing SEAP expression. Co-transfection of a vector constitutively expressing mUTS (pCK25) and the vector with SEAP controlled by P_{UREX8} (pCK9) into human cervical adenocarcinoma cells (HeLa), human embryonic kidney cells (HEK-293) and human fibrosarcoma cells (HT-1080) grown in medium free of uric acid showed that mUTS bound and silenced P_{UREX8}-driven SEAP expression up to 98% (**Fig. 1c**), reaching a tightness which was previously only achieved by more complex transcription-translation networks[16]. Constitutive expression of mUTS was driven by the human elongation factor 1α promoter ($P_{hEF1\alpha}$). Addition of 5 mM uric acid to transfected cultures triggered the release of mUTS from $hucO_8$ and derepressed SEAP expression (**Fig. 1d–f**).

To increase the sensitivity of the UREX circuit to uric acid, we co-transfected it into cells with a third plasmid that expressed the human urate-anion transporter URAT1, which is naturally involved in renal urate clearance and maintenance of uric acid homeostasis[17,18]. The transporter's urate-uptake capacity is expected to increase intracellular uric acid levels and consequently amplify UREX sensitivity. Transfecting a constitutive URAT1 expression vector (pURAT1) into HeLa, HEK-293 and HT-1080 cells engineered for UREX-controlled SEAP expression resulted in considerably higher SEAP levels (**Fig. 1d–f**) compared with URAT1-free cells or with cells that constitutively expressed a control anion transporter OAT2 (human organic anion transporter 2), which is unrelated to uric acid transport. The expression of either endogenous or transgenic URAT1 and OAT2 was determined by semiquantitative RT-PCR and demonstrated that neither HeLa, HEK-293 nor HT-1080 wild-type cells express endogenous URAT1 or OAT2 (**Supplementary Fig. 1** and **Supplementary Table 2**).

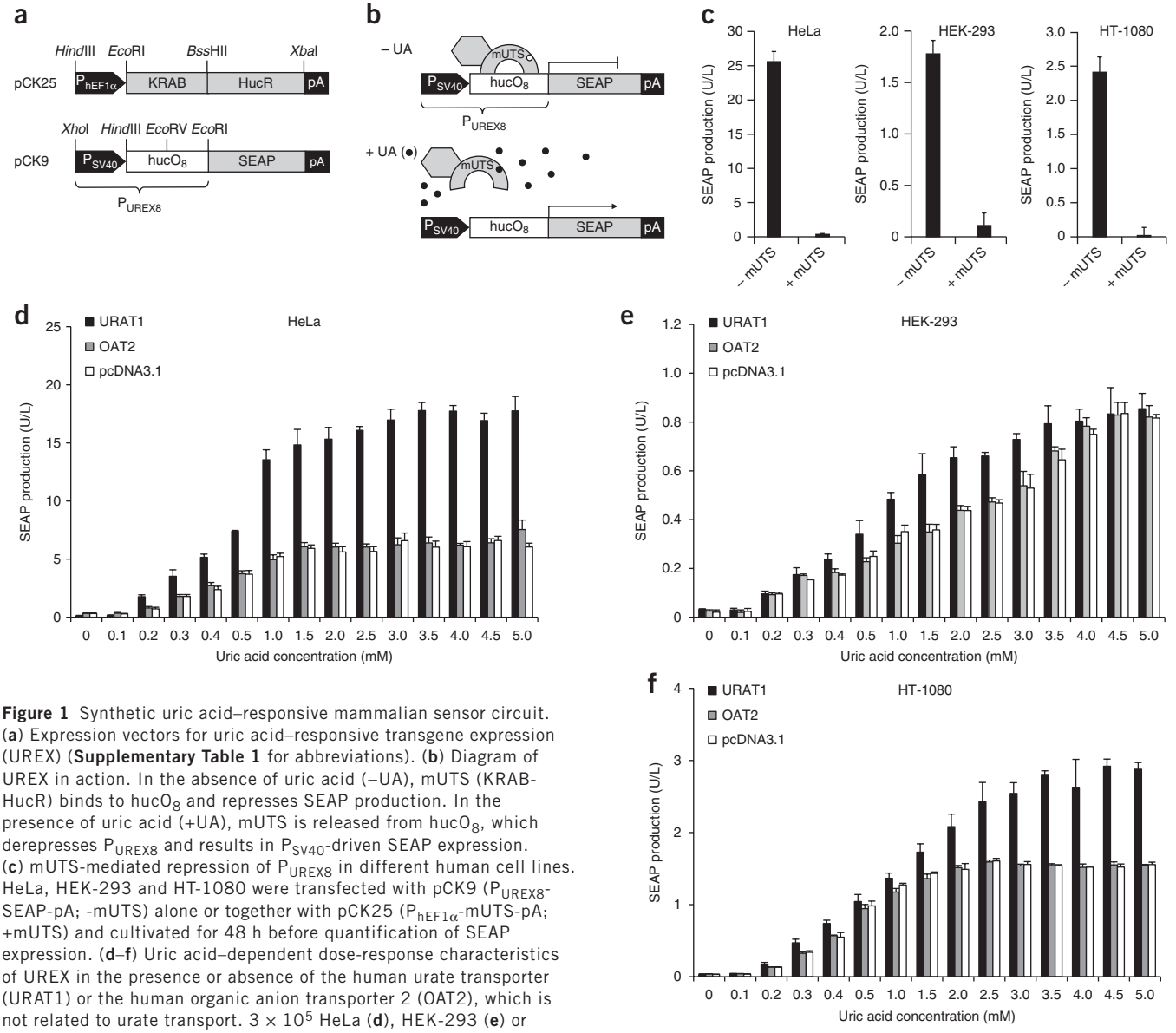


Figure 1 Synthetic uric acid–responsive mammalian sensor circuit. (**a**) Expression vectors for uric acid–responsive transgene expression (UREX) (**Supplementary Table 1** for abbreviations). (**b**) Diagram of UREX in action. In the absence of uric acid (–UA), mUTS (KRAB-HucR) binds to $hucO_8$ and represses SEAP production. In the presence of uric acid (+UA), mUTS is released from $hucO_8$, which derepresses P_{UREX8} and results in P_{SV40}-driven SEAP expression. (**c**) mUTS-mediated repression of P_{UREX8} in different human cell lines. HeLa, HEK-293 and HT-1080 were transfected with pCK9 (P_{UREX8}-SEAP-pA; -mUTS) alone or together with pCK25 ($P_{hEF1\alpha}$-mUTS-pA; +mUTS) and cultivated for 48 h before quantification of SEAP expression. (**d–f**) Uric acid–dependent dose-response characteristics of UREX in the presence or absence of the human urate transporter (URAT1) or the human organic anion transporter 2 (OAT2), which is not related to urate transport. 3×10^5 HeLa (**d**), HEK-293 (**e**) or HT-1080 (**f**) were co-transfected with the UREX sensor plasmids pCK9 and pCK25 (**Fig. 1a**) and either pURAT1 (P_{hCMV}-URAT1-pA), pOAT2 (P_{hCMV}-OAT2-pA) or the isogenic control vector pcDNA3.1. The transfected populations were cultivated in medium supplemented with different uric acid concentrations and resulting SEAP levels were profiled after 48 h.

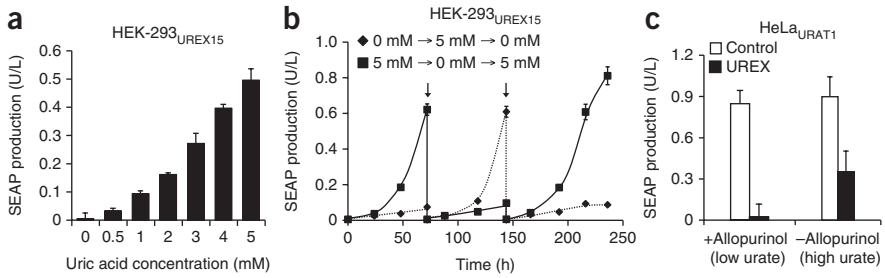


Figure 2 Validation of UREX-controlled SEAP expression in transgenic HEK-293 and urate oxidase–deficient mice. (**a**) Uric acid–based dose response profile of the triple-transgenic HEK-293$_{UREX15}$ cell line stably engineered for UREX-controlled SEAP and constitutive URAT1 expression. 5×10^4 HEK-293$_{UREX15}$ cells were cultivated in the presence of different uric acid concentrations and SEAP levels were quantified in the culture supernatant after 72 h. (**b**) Reversibility of the UREX-based uric acid sensor circuit. 2×10^5 HEK-293$_{UREX15}$ cells were cultivated for 10 d while uric acid concentrations were alternated from 0–5 mM every 72 h (arrows). (**c**) HeLa$_{URAT1}$ cells engineered for UREX-controlled SEAP expression were microencapsulated in coherent alginate-poly-L-lysine-alginate microcapsules and intraperitoneally injected (2×10^6 cells per mouse, 200 cells/capsule) into *uox*$^{-/-}$ mice that had received 150 μg/ml (wt/vol) of the hyperuricemia therapeutic allopurinol (+allopurinol) in their drinking water to reduce urate levels (UREX) or untreated *uox*$^{-/-}$ mice exhibiting pathologic urate levels (–allopurinol). Control implants contained cells transgenic for constitutive SEAP expression (control). SEAP levels were profiled in the serum of the animals 72 h after cell implantation.

HeLa cells were particularly more sensitive within the urate range typically found in the human blood (200–500 μM) (**Fig. 1d–f**).

To characterize adjustability and long-term reversibility of the mammalian uric acid sensor system, we constructed the stable human cell line HEK-293$_{UREX15}$ that is triple-transgenic for UREX-controlled SEAP and constitutive URAT1 expression. SEAP expression of HEK-293$_{UREX15}$ could be precisely adjusted and showed a progressive increase in response to escalating concentrations of uric acid (**Fig. 2a**). Reversibility of UREX-controlled SEAP production was assessed by cultivating HEK-293$_{UREX15}$ for 10 d while alternating uric acid concentrations from 0 mM to 5 mM every 72 h. UREX control was completely reversible and could be reset at any time without showing any expression memory effect on SEAP production (**Fig. 2b**). Indeed, 72 h after intraperitoneal implantation of microencapsulated cells engineered for URAT1 and UREX-controlled SEAP expression into urate oxidase-deficient (*uox*$^{-/-}$) mice that were developing hyperuricemia with human-like symptoms[19], the synthetic urate sensor circuit was sufficiently sensitive to discriminate between mice developing hyperuricemia (high urate levels (23.1 ± 5.6 (s.d.) mg/dl); 0.35 ± 0.15 (s.d.) U/L SEAP) and animals that received the licensed urate-reducing arthritis therapeutic allopurinol (Zyloprim) in their drinking water (low urate levels (7.8 ± 3.1 (s.d.) mg/dl); 0.02 ± 0.10 (s.d.) U/L SEAP) (**Fig. 2c**). Control mice treated with cells constitutively producing SEAP showed unaltered SEAP expression in the presence (low urate levels; 0.85 ± 0.10 (s.d.) U/L SEAP) or absence (high urate levels; 0.90 ± 0.15 (s.d.) U/L SEAP) of allopurinol. Taken together, these results suggest that the precise and robust uric acid sensor UREX may be suitable for therapeutic control of this pathologic metabolite.

For feedback-controlled reduction of uric acid levels in an autonomous and self-sufficient manner, the uric acid sensor circuit must be linked to expression of a uricase/urate oxidase (Uox) which converts urate to the more soluble and renally secretable allantoin[20]. We therefore optimized the codons of the cofactor-independent *Aspergillus flavus* uricase for expression in mammalian cells obtaining mUox.

mUox reduced uric acid (0.5 mM) in the culture medium when constitutively expressed (P$_{UREX8}$-mUox-pA, referred to as pCK67) or controlled by mUTS (pCK67 with pCK25, which constitutively expresses mUTS) (**Fig. 3a**). Urate reduction was more efficient when mUox was engineered for mammalian cell–based secretion by in-frame fusion to an immunoglobulin-derived secretion signal[21] (SS$_{Igk}$). This optimized uricase, which is referred to as smUox, was used in further studies and was expressed on plasmid pCK65 (P$_{UREX8}$-SS$_{Igk}$-mUox-pA).

The dynamics of UREX-controlled uric acid metabolism were profiled during a 120-h cultivation of HeLa$_{URAT1}$ (a HeLa-derived cell line transgenic for constitutive URAT1 expression) engineered for constitutive mUTS and P$_{UREX8}$-driven smUox expression (**Fig. 3b**). Isogenic HeLa$_{URAT1}$ control cells that constitutively expressed smUox (without mUTS co-expression) cleared uric acid from the medium. In contrast, the uric acid levels of cell cultures harboring the synthetic UREX-smUox control circuit leveled out at 0.14 mM (a concentration suggested to support oxidative-stress protection in a physiologic context), even when exposed to different starting uric acid concentrations. These results indicate that uric acid levels have fallen below a threshold concentration that was no longer able to induce UREX control and trigger further urate oxidation (**Fig. 3b**). The uric acid homeostasis level of 0.14 mM was consistently reached even when the UREX-based control device was challenged with a second dose of uric acid after 72 h (**Fig. 3c**). Overall, these data confirm self-sufficient feedback control of uric acid levels by UREX-smUox and suggest that this synthetic circuit could reduce pathologic urate levels into a physiologic window in an auto-controlled manner.

To further elucidate how the UREX system and its key design features, such as the number of hucO sites in P$_{UREX}$, would influence uric acid homeostasis *in vivo*, we established a dynamic mathematical model based on ordinary differential equations. Briefly, we developed an eight-state model that describes all relevant processes such as (controlled) gene expression, transport and other enzyme-catalyzed reactions. Model parameters were constrained by literature data and further parameter estimation using part of the experimental data for the UREX system in HeLa cells. We validated the model with independent predictions, and obtained good matches between simulation results and experimental data in both scenarios. In addition, changes in a subset of cell line–specific model parameters allowed us to quantitatively capture UREX dose responses in HEK-293 and HT-1080 cells as well (**Supplementary Results**, **Supplementary Figs. 2–7** and **Supplementary Tables 3** and **4**).

To predict the behavior after encapsulation of the HeLa$_{URAT1}$-based system in mice, we initially used a simplified mathematical model where the implant and the cells inside are in a flow with constant influx and identical outflux with respect to the mouse body, such that the liquid volume is constant. This situation corresponds to the setup of a chemostat with cell retention for different constant dilution rates and input uric acid concentrations. The model predictions show that, for large ranges of these parameters, the system will achieve a desirable steady-state uric acid output concentration in the physiological range. Moreover, with eight HucR binding sites per P$_{UREX}$, the system will be rather insensitive to variations in uric acid

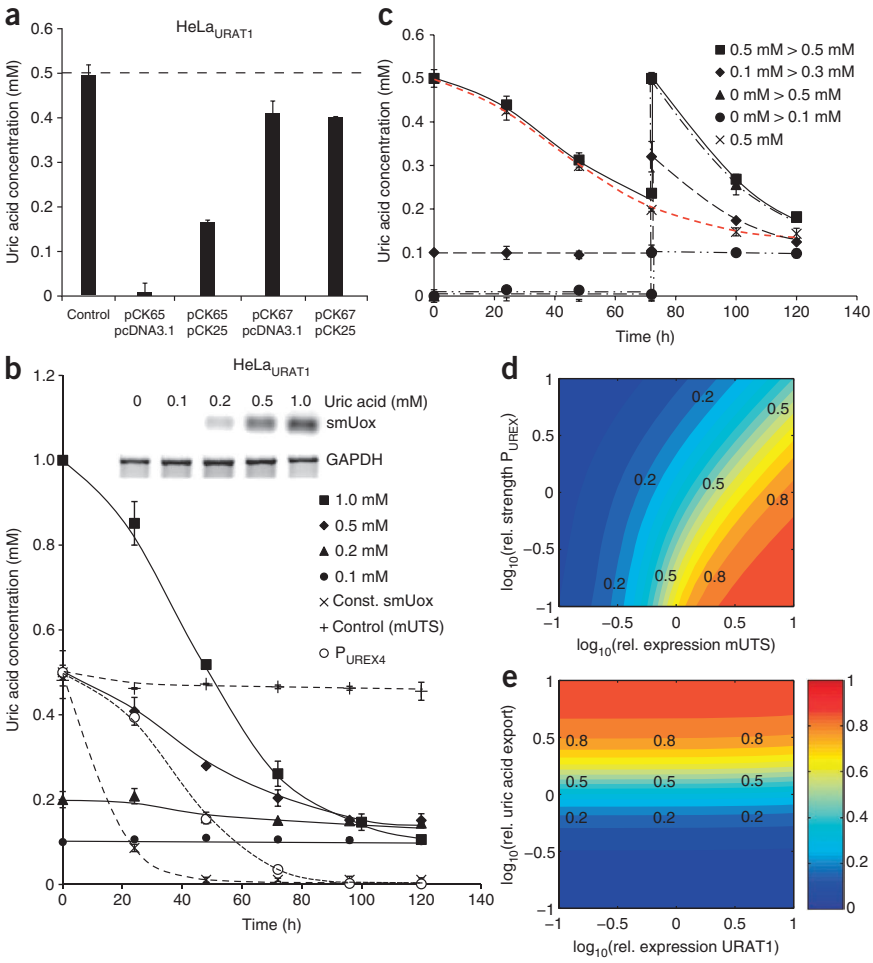


Figure 3 Functional characterization of an engineered mammalian *A. flavus*–derived urate oxidase. (**a**) Profiling of urate reduction mediated by constitutive or UREX-controlled expression of an intracellular (mUox) or secretion-engineered (smUox) urate oxidase variant. 2×10^5 HeLa$_{URAT1}$ cells were co-transfected with either pCK65 and isogenic control vector pcDNA3.1, pCK65 and pCK25, pCK67 and pcDNA3.1, or pCK67 and pCK25. The cells were cultivated for 72 h with 0.5 mM uric acid before uric acid levels were determined in the culture supernatant. (**b**) Uric acid degradation profiles of UREX-controlled smUox expression. 2×10^5 HeLa$_{URAT1}$ cells engineered for UREX-controlled smUox expression (solid lines) were cultivated in medium containing starting uric acid concentrations of either 0.1, 0.2, 0.5 or 1 mM and uric acid reduction kinetics were profiled for 120 h. Control experiments show 2×10^5 HeLa$_{URAT1}$ engineered for constitutive smUox or mUTS expression (dashed lines). The upper panel shows smUox and GAPDH (control) transcript levels profiled 24 h after induction. (**c**) Uric acid homeostasis. 2×10^5 HeLa$_{URAT1}$ cells engineered for UREX-controlled smUox expression were cultivated in medium containing starting concentrations of 0, 0.1 or 0.5 mM uric acid. After 72 h the cultures received a second dose of uric acid (0.1, 0.3, 0.5 mM) after which the UREX-based control device again adjusted uric acid concentrations to the homeostasis level (dashed red line, 0.5 mM uric acid profile of **Fig. 3b**). (**d**,**e**) Model predictions for uric acid homeostasis using a physiological model. Steady-state values of uric acid concentration obtained by simulating the interplay of UREX circuit in HeLa$_{URAT1}$ cells and of body cells connected by a fluid flow (**Supplementary Results**). Contour plots of steady-state extracellular urate concentration as a function of relative mUTS expression and of relative P$_{UREX}$ promoter strength (**d**), and as a function of relative URAT1 expression and of maximal uric acid export rate (**e**), respectively. Colors indicate concentrations according to the color bar in **e**; see also labels, values in mM.

input, but substantially fewer operator sites will result in unphysiologically low levels because of too-leaky promoter characteristics (**Supplementary Results** and **Supplementary Fig. 8**).

We tested these predictions by developing a more detailed promoter model and by experimentally constructing UREX circuits with P$_{UREX}$ variants harboring fewer than eight binding sites. Both approaches showed very good quantitative agreement with the earlier predictions. An example for UREX-controlled dynamics with P$_{UREX4}$ is shown in **Figure 3b** (**Supplementary Results**, **Supplementary Figs. 9–12** and **Supplementary Table 5**). Finally, we embedded the model into a more realistic physiological setting by conceptually splitting the mouse into two compartments, one containing the UREX implant and the other one consisting of the rest of the animal, respectively. For broad ranges of parameter values this model confirmed uric acid homeostasis (**Supplementary Results** and **Supplementary Fig. 13**). In addition, it allowed for detailed predictions of possibilities for fine-tuning the circuit. For instance, the model predicts that variation of mUTS levels is more effective in adjusting the circuit than replacing the promoter in P$_{UREX}$ (**Fig. 3d**). Surprisingly, the model reveals that manipulation of uric acid export is more promising than adjustment of URAT1 expression, to which the system is not sensitive at all (**Fig. 3e**). Hence, the model predictions further reinforce the potential use of the synthetic circuit for uric acid homeostasis in pathological conditions, and they suggest key design features of the circuit.

When urate oxidase–deficient *uox*$^{-/-}$ mice were treated with cell implants expressing UREX-controlled smUox, urate levels in serum and in urine dropped to concentrations reached by standard allopurinol therapy. In contrast, control animals (e.g., HeLa$_{URAT1}$ cell implants, no allopurinol) accumulated urate and developed acute hyperuricemia in the absence of allopurinol with acute uric acid crystal deposition in renal tubules resembling those of the tumor lysis syndrome in humans[19] (**Fig. 4a**,**b**). Histologic analysis of H&E-stained kidney sections from all treatment groups confirmed that UREX-smUox-based therapeutic cell implants were able to considerably reduce uric acid crystal deposits in the proximal tubules (**Fig. 4c–f**).

Gout is the most common inflammatory arthritis, affecting >1% of the human population in industrialized countries. Clinical manifestations, which result from monosodium urate crystal deposition, include acute gouty arthritis, chronic gouty arthropathy, tophi, renal dysfunction and urolithiasis[9,11]. Owing to a urate oxidase deficiency, urate blood levels of humans are up to 50 times higher than in other mammals, which increases the risk of developing hyperuricemic diseases[11,17]. Uric acid has been suggested to be a free radical scavenger that may protect living systems from lipid peroxidation, DNA

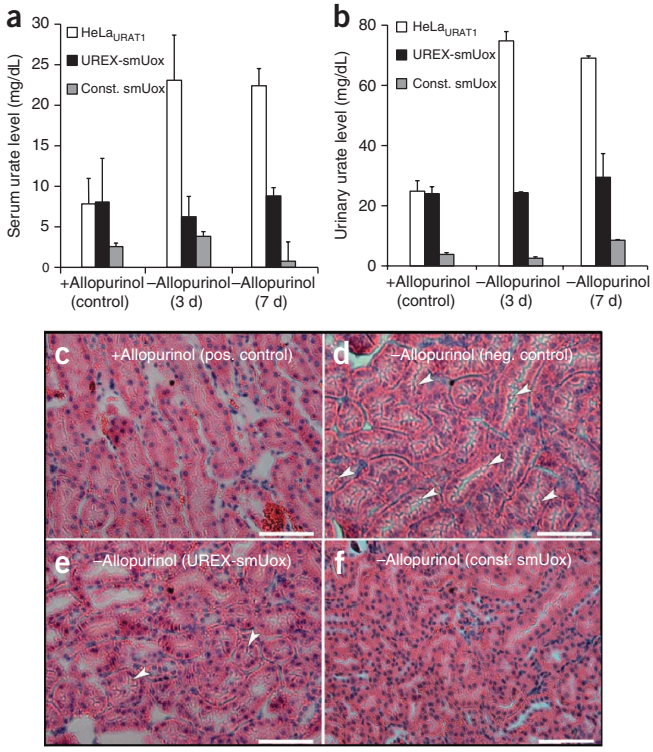


Figure 4 UREX-controlled smUox-mediated reduction of pathologic urate levels in mice. Microencapsulated HeLa$_{URAT1}$ cells engineered for UREX-controlled smUox expression were intraperitoneally implanted (2×10^6 cells per mouse) into untreated *uox*$^{-/-}$ mice exhibiting pathologic urate levels or into *uox*$^{-/-}$ mice having received 150 µg/ml (wt/vol) of the hyperuricemia therapeutic allopurinol in their drinking water (UREX-smUox) (control). Control implants contained parental HeLa$_{URAT1}$. (**a**,**b**) Urate levels were profiled in serum (**a**) and urine (collected for 24 h) (**b**) of the animals 3 and 7 d after cell implantation. (**c**–**f**) Tissue sections showing anisotropic uric acid crystals (arrow) in the kidneys of *uox*$^{-/-}$ mice receiving 150 µg/ml (wt/vol) allopurinol (positive control) (**c**), phosphate-buffered saline (negative control) (**d**), implants with HeLa$_{URAT1}$ engineered for UREX-controlled smUox expression (**e**) and implants with HeLa$_{URAT1}$ engineered for constitutive smUox expression (**f**). Quantitative morphometric analysis revealed 3.3% ± 2.9 (s.d.)/1.5% ± 0.8 (s.d.) (**c**), 109.1% ± 27.3 (s.d.)/26.9% ± 5.6 (s.d.) (**d**), 10.4% ± 4.6(s.d.)/3.7% ± 1.5 (s.d.) (**e**) and 0.4% ± 0.5 (s.d.)/0.6% ± 0.7 (s.d.) (**f**), percent crystals/percent relative area of crystalline deposits per tubulus and/or tubulus lumen profile of the respective treatment group. Scale bars, 100 µm.

and protein damage, mitochondrial dysfunction[22,23] and diseases such as cancer[24] and neurodegenerative disorders[22,25,26]. Uric acid should therefore be carefully regulated to balance its positive and negative physiologic effects[25].

Two urate-reducing drug classes have been developed for the treatment of chronic gouty arthritis[9]. Uricosuric drugs such as probenecid (Probenecid), losartan (Hyzaar) and benzbromarone inhibit URAT1, thereby preventing reabsorption and increasing uric acid excretion. Xanthine oxidase inhibitors, such as allopurinol, oxypurinol (Oxyprim) and Febuxostat (Uloric)[27], which has recently been licensed by the US Food and Drug Administration for the treatment of chronic hyperuricemia, block metabolic production of uric acid. Xanthine oxidase inhibitors are used as first-line treatment of patients with renal calculi, renal insufficiency, urate overproduction and patients receiving simultaneous diuretic and cyclosporin therapies. In contrast, uricosuric drugs provide the treatment of choice in allopurinol-allergic patients and underexcretors with normal renal function and no history of urolithiasis. Cell implants harboring a UREX-based network are prosthetic devices which provide self-sufficient and reversible control of uric acid levels in the bloodstream by preventing critical urate accumulation while also preserving basal uric acid levels that may improve urate-based protection from oxidative stress. Such devices may therefore be equally suited as a preventive treatment and as a therapy for acute hyperuricemic disorders including the tumor lysis syndrome and gout.

By functionally assembling a human urate transporter, a hucO-HucR–derived sensor circuit and sensor-controlled expression of a secretion-engineered version of a clinically licensed *A. flavus* urate oxidase (or uricase) we have designed a synthetic mammalian uric acid homeostasis network that enables (i) monitoring of urate levels in the blood, (ii) automatic induction of the core sensor unit at pathologic urate levels, (iii) prompt reduction of urate by sensor-controlled expression of clinically licensed urate oxidase and (iv) spontaneous shutdown when physiologic urate levels have been reached. UREX-based control of urate levels in urate oxidase–deficient mice, which develop hyperuricemia with human-like symptoms[19], showed that the synthetic circuit operated as expected and most importantly within the clinically relevant concentration range.

In humans, formation of monosodium urate crystals leading to painful inflammations typically occurs at blood urate levels of above 6.8 mg/dl[11,23,24]. The crystals are known to dissolve at a circulation level below 6 mg/dl[9], and cure of chronic hyperuricemia would require long-term maintenance of serum urate levels below this level. The UREX sensor is able to sense such pathologic levels and trigger dose-dependent expression of secreted uricase, which consequently reduces urate in the bloodstream of treated animals to 5 mg/dl.

The simple design principle of this therapeutic circuit may motivate the assembly of other prosthetic networks that sense metabolic disturbances and circulating pathologic metabolites, process such signals and coordinate an adjusted therapeutic response. Provided that prosthetic networks are robust and operate in safely contained cell implants that can be readily replaced at reasonable intervals, such technology might provide standard preventive surveillance as well as precise interventions in acute situations.

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 W. Weber (Albert-Ludwigs-Universität Freiburg, Germany) for conceptual input and for providing pBluescript-mHucR, M. Gilet for skillful assistance with the animal study and Marcia Schoenberg for critical comments on the manuscript. This work was supported by the Swiss National Science Foundation (grant no. 31003A-126022) and in part by the EC Framework 7 (Persist).

AUTHOR CONTRIBUTIONS

C.K., J.S. and M.F. designed the project, analyzed results and wrote the manuscript. C.K., M.G., M.D.-E.B. and V.D. performed the experimental work. J.S. designed the mathematical model and performed simulations.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

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

1. Kramer, B.P. & Fussenegger, M. Hysteresis in a synthetic mammalian gene network. *Proc. Natl. Acad. Sci. USA* **102**, 9517–9522 (2005).
2. Kramer, B.P. *et al.* An engineered epigenetic transgene switch in mammalian cells. *Nat. Biotechnol.* **22**, 867–870 (2004).
3. Fussenegger, M. Synthetic biology: Synchronized bacterial clocks. *Nature* **463**, 301–302 (2010).
4. Tigges, M., Marquez-Lago, T.T., Stelling, J. & Fussenegger, M. A tunable synthetic mammalian oscillator. *Nature* **457**, 309–312 (2009).
5. Rinaudo, K. *et al.* A universal RNAi-based logic evaluator that operates in mammalian cells. *Nat. Biotechnol.* **25**, 795–801 (2007).
6. Weber, W., Daoud-El Baba, M. & Fussenegger, M. Synthetic ecosystems based on airborne inter- and intrakingdom communication. *Proc. Natl. Acad. Sci. USA* **104**, 10435–10440 (2007).
7. Dean, J.T. *et al.* Resistance to diet-induced obesity in mice with synthetic glyoxylate shunt. *Cell Metab.* **9**, 525–536 (2009).
8. Cairo, M.S. & Bishop, M. Tumour lysis syndrome: new therapeutic strategies and classification. *Br. J. Haematol.* **127**, 3–11 (2004).
9. Terkeltaub, R.A. Clinical practice. Gout. *N. Engl. J. Med.* **349**, 1647–1655 (2003).
10. Liebman, S.E., Taylor, J.G. & Bushinsky, D.A. Uric acid nephrolithiasis. *Curr. Rheumatol. Rep.* **9**, 251–257 (2007).
11. Terkeltaub, R., Bushinsky, D.A. & Becker, M.A. Recent developments in our understanding of the renal basis of hyperuricemia and the development of novel antihyperuricemic therapeutics. *Arthritis Res. Ther.* **8**, S4 (2006).
12. Daly, M.J. A new perspective on radiation resistance based on *Deinococcus radiodurans*. *Nat. Rev. Microbiol.* **7**, 237–245 (2009).
13. Wilkinson, S.P. & Grove, A. HucR, a novel uric acid-responsive member of the MarR family of transcriptional regulators from *Deinococcus radiodurans*. *J. Biol. Chem.* **279**, 51442–51450 (2004).
14. Wilkinson, S.P. & Grove, A. Negative cooperativity of uric acid binding to the transcriptional regulator HucR from *Deinococcus radiodurans*. *J. Mol. Biol.* **350**, 617–630 (2005).
15. Bellefroid, E.J., Poncelet, D.A., Lecocq, P.J., Revelant, O. & Martial, J.A. The evolutionarily conserved Kruppel-associated box domain defines a subfamily of eukaryotic multifingered proteins. *Proc. Natl. Acad. Sci. USA* **88**, 3608–3612 (1991).
16. Deans, T.L., Cantor, C.R. & Collins, J.J. A tunable genetic switch based on RNAi and repressor proteins for regulating gene expression in mammalian cells. *Cell* **130**, 363–372 (2007).
17. Oda, M., Satta, Y., Takenaka, O. & Takahata, N. Loss of urate oxidase activity in hominoids and its evolutionary implications. *Mol. Biol. Evol.* **19**, 640–653 (2002).
18. Enomoto, A. *et al.* Molecular identification of a renal urate anion exchanger that regulates blood urate levels. *Nature* **417**, 447–452 (2002).
19. Wu, X. *et al.* Hyperuricemia and urate nephropathy in urate oxidase-deficient mice. *Proc. Natl. Acad. Sci. USA* **91**, 742–746 (1994).
20. Legoux, R. *et al.* Cloning and expression in *Escherichia coli* of the gene encoding Aspergillus flavus urate oxidase. *J. Biol. Chem.* **267**, 8565–8570 (1992).
21. Fluri, D.A., Kemmer, C., Daoud-El Baba, M. & Fussenegger, M. A novel system for trigger-controlled drug release from polymer capsules. *J. Control. Release* **131**, 211–219 (2008).
22. Pacher, P., Beckman, J.S. & Liaudet, L. Nitric oxide and peroxynitrite in health and disease. *Physiol. Rev.* **87**, 315–424 (2007).
23. Whiteman, M. & Halliwell, B. Protection against peroxynitrite-dependent tyrosine nitration and alpha 1-antiproteinase inactivation by ascorbic acid. A comparison with other biological antioxidants. *Free Radic. Res.* **25**, 275–283 (1996).
24. Ames, B.N., Cathcart, R., Schwiers, E. & Hochstein, P. Uric acid provides an antioxidant defense in humans against oxidant- and radical-caused aging and cancer: a hypothesis. *Proc. Natl. Acad. Sci. USA* **78**, 6858–6862 (1981).
25. Kutzing, M.K. & Firestein, B.L. Altered uric acid levels and disease states. *J. Pharmacol. Exp. Ther.* **324**, 1–7 (2008).
26. Schwarzschild, M.A. *et al.* Serum urate as a predictor of clinical and radiographic progression in Parkinson disease. *Arch. Neurol.* **65**, 716–723 (2008).
27. Schumacher, H.R. Jr., Becker, M.A., Lloyd, E., MacDonald, P.A. & Lademacher, C. Febuxostat in the treatment of gout: 5-yr findings of the FOCUS efficacy and safety study. *Rheumatology (Oxford)* **48**, 188–194 (2009).

ONLINE METHODS

Vector construction. Design details of all plasmids and oligonucleotides are provided in **Supplementary Table 1**.

Cell culture and transfections. HeLa (American Type Culture Collection (ATCC)), HEK-293 (ATCC) and HT-1080 (ATCC) were cultivated in DMEM (Invitrogen) supplemented with 10% (vol/vol) FCS (PAN Biotech) and 1% (vol/vol) penicillin/streptomycin solution (P/S, PAN Biotech). All cell types were cultivated at 37 °C in a humidified atmosphere containing 5% CO_2. HEK-293 were transfected using a standard $CaPO_4$-based protocol[2]. HeLa were also transfected according to this standard $CaPO_4$-based protocol, with the exception that the DNA precipitates were incubated with the cells for 12 h before changing the medium. HT-1080 were transfected with FuGENE 6 (Roche Diagnostics) according to the supplier's procedure. Transfected cells were cultivated in DMEM supplemented with 10% (vol/vol) knockout serum replacement (KOSR, Invitrogen), 1% P/S and, optionally, with uric acid (Acros Organics) (standard medium). Transgene expression values were normalized for variations in transfection efficiency by parallel transfections using the constitutive EYFP-expression vector pDF60. To establish uric acid-dependent dose-response characteristics, 3×10^5 cells were seeded per well of a six-well plate (Thermo Fisher Scientific) and transfected as described above. The cells were trypsinized (200 µl trypsin; PAN Biotech; 5 min, 37 °C), collected by centrifugation (2 min, 120*g*) and resuspended in 1.5 ml standard medium. We transferred 100 µl of this cell suspension to individual wells of a 96-well plate, supplemented cells with different concentrations of uric acid and cultivated them for 48 h before reporter gene expression was profiled in the culture supernatant.

Semiquantitative RT-PCR. 1 µg of total cellular RNA, isolated from mock-, pURAT1-, pOAT2-, or pCK65 and pCK25 transfected HeLaURAT1, HEK-293 and HT-1080 cells using the RNeasy Mini Kit (Qiagen), was reverse transcribed using the TaqMan Reverse Transcription Reagents (Applied Biosystems) and the RNA levels were quantified with the Mastercycler ep gradient/S system (Vaudaux-Eppendorf) and probes specific for URAT1, OAT2, smUox and the glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (**Supplementary Table 2**).

Analytical assays. SEAP levels were quantified in cell culture supernatants and mouse serum using a standard *p*-nitrophenylphosphate-based light-absorbance time course[2,6]. Uric acid levels were assessed in cell culture supernatants, murine serum or urine using the Amplex Red uric acid/uricase assay kit (Invitrogen) according to the manufacturer's protocol. In brief, uric acid–containing samples were diluted in reaction buffer. The addition of uricase triggered the enzymatic conversion of uric acid to allantoin, CO_2 and H_2O_2, which reacts, in the presence of horseradish peroxidase stoichiometrically with Amplex Red reagent to generate the red-fluorescent oxidation product resorufin, which can be quantified at 585 nm.

Construction of stable transgenic cell lines. HeLa-derived $HeLa_{URAT1}$ cells, transgenic for the constitutive expression of the human renal urate-anion exchanger (URAT1), were constructed by co-transfection of pURAT1 (P_{hCMV}-URAT1-pA; ImaGenes) and pZeoSV2 (conferring resistance to zeocin; Invitrogen) at a ratio of 10:1. After a 14-d selection period using 200 µg/ml (wt/vol) zeocin cells were clonally expanded and individual clones were profiled for URAT1 expression using qRT-PCR. The cell line $HeLa_{URAT1}$ was chosen for further experiments. The triple-transgenic HEK-293–derived $HEK\text{-}293_{UREX15}$ cell line, enabling urate-inducible SEAP expression, was constructed by sequential co-transfection and clonal selection of (i) pCK9 (P_{UREX8}-SEAP-pA) and pCK25 ($P_{hEF1\alpha}$-KRAB-mHucR-pA, also carrying a constitutive expression cassette conferring resistance to blasticidin) (ratio of 1:1, 14-d selection in DMEM containing 10% FCS and 20 µg/ml (wt/vol) blasticidin (InvivoGen)) and (ii) pURAT1 (P_{hCMV}-URAT1-pA) and pZeoSV2 (conferring resistance to zeocin) (ratio of 10:1, 14-d selection in DMEM containing 10% FCS, 20 µg/ml (wt/vol) blasticidin and 200 µg/ml (wt/vol) zeocin (InvivoGen). Individual $HEK\text{-}293_{UREX}$ clones were randomly picked and assessed for urate-triggered SEAP expression. $HEK\text{-}293_{UREX15}$ was chosen for further studies.

***In vivo* methods.** Urate oxidase-deficient mice ($uox^{-/-}$) developing hyperuricemia with human-like symptoms[19] were used to validate the synthetic UREX-based uric acid control network *in vivo* (Charles River Laboratories). Because $uox^{-/-}$ mice die within 4 weeks of age, breeding and long-term maintenance requires the addition of 150 µg/ml (wt/vol) allopurinol (Sigma) to the drinking water[19]. Two weeks before implantation of UREX-transgenic cells the allopurinol treatment of urate oxidase–deficient mice was either continued or stopped to produce animal groups with low or high pathogenic uric acid levels in the bloodstream, respectively. $HeLa_{URAT1}$ mice, transgenic for constitutive URAT1 expression, were either transiently transfected with pCK65 to provide an isogenic constitutive smUox expression control or co-transfected with pCK25 and either pCK9, pSEAP2-control or pCK65 and then microencapsulated in 200 µm alginate-(poly-L-lysine)-alginate capsules (200 cells/capsule) using an Inotech Encapsulator Research IE-50R (EncapBioSystems), according to the manufacturer's protocol and applying the following settings: 200 µm single nozzle, stirrer speed control set to 5 units, 20-ml syringe with a flow rate of 410 units, nozzle vibration frequency 1,024 Hz, voltage for capsule dispersion 900 V. 700 µl PBS containing 2×10^6 encapsulated cells (10^4 capsules/mouse) was intraperitoneally injected into urate oxidase–deficient mice. Control mice were implanted with microencapsulated parental $HeLa_{URAT1}$. Two and 6 d after implantation, the mice were transferred to a clean cage and the urine of each treatment group was sampled for 24 h. Three and 7 d after implantation the mice were euthanized, their blood was collected and the serum isolated in microtainer SST tubes (Becton Dickinson) according to the manufacturer's protocol. All experiments involving animals were performed according to the directive of the European Community Council (86/609/EEC), approved by the French Republic (no. 69266310) and carried out by M.D.E. at the Institut Universitaire de Technologie, IUTA, F-69622 Villeurbanne Cedex, France.

Mathematical modeling. All details on model development are provided in the **Supplementary Results**, **Supplementary Figures 2–13** and **Supplementary Tables 3–6**.

Histology and quantification of uric acid crystals in renal tubules. For microscopic analysis of uric acid crystal deposits in the kidney of $uox^{-/-}$ mice implanted with microencapsulated cells transgenic for constitutive URAT1 and UREX-controlled smUox expression–treated animals were split into treatment groups receiving 150 µg/ml (wt/vol) or no allopurinol in their drinking water as described above. Seven days after implantation, the mice were euthanized, their kidneys explanted and fixed in 4% (wt/vol) paraformaldehyde (Sigma) in PBS for 4 h. The kidneys were washed in 10% (wt/vol) sucrose (Sigma) in PBS for 90 min, trimmed, dehydrated and embedded in paraffin. An ultracut device (Zeiss) was used to obtain 3-µm sections of the tissues, which were transferred to gelatinized microslides and air-dried overnight at 37 °C. After paraffin removal by submersion in xylene (3×10 min), the tissues were rehydrated by sequential incubation (10 min) in decreasing ethanol concentrations (90%, 80%, 70%), rinsed twice in TBS (Tris-buffered saline; 50 mM Tris/HCl (pH 7.4), 100 mM NaCl) and stained with H&E solution. The samples were visualized under polarized light to assess anisotropism using a Leica DMRBE microscope. The number of crystals as well as the total crystalline area of kidney sections, derived from at least five animals per treatment group, was quantified using the iTEM morphometry software (Olympus Soft Imaging Solutions).

nature biotechnology

Derivation, propagation and controlled differentiation of human embryonic stem cells in suspension

Debora Steiner[1], Hanita Khaner[1], Malkiel Cohen[1], Sharona Even-Ram[1], Yaniv Gil[1], Pavel Itsykson[1], Tikva Turetsky[1], Maria Idelson[1], Einat Aizenman[2], Rita Ram[3], Yael Berman-Zaken[1] & Benjamin Reubinoff[1]

Undifferentiated human embryonic stem cells (hESCs) are currently propagated on a relatively small scale as monolayer colonies[1–7]. Culture of hESCs as floating aggregates is widely used for induction of differentiation into embryoid bodies[8]. Here we show that hESC lines can be derived from floating inner cell masses in suspension culture conditions that do not involve feeder cells or microcarriers. This culture system supports prolonged propagation of the pluripotent stem cells as floating clusters without their differentiation into embryoid bodies. HESCs cultivated as aggregates in suspension maintain the expression of pluripotency markers and can differentiate into progeny of the three germ layers both *in vitro* and *in vivo*. We further show the controlled differentiation of hESC clusters in suspension into neural spheres. These results pave the way for large-scale expansion and controlled differentiation of hESCs in suspension, which would be valuable in basic and applied research.

HESCs hold great promise as a renewable source of cells for basic and applied research[1–7]. hESC lines have so far been derived in monolayer cultures from blastocysts, inner cell masses or single blastomeres that were plated on supporting cell layers[9]. In one instance, two hESC lines were derived in a feeder-free monolayer system on extracellular matrix (ECM) proteins[3]. hESC lines are most commonly expanded as adherent colonies, whereas the current notion is that detachment into free-floating clusters induces differentiation. However, adherent culture is a major limitation for large-scale expansion of the cells, that could be overcome by propagation in suspension. Recently, hESCs were expanded either attached to coated microcarriers or unattached in spinner flasks for short periods[10–13]. Here we show the derivation, prolonged propagation and controlled differentiation of hESCs in a defined microcarrier-free suspension culture system. Performing the whole process in a defined suspension system sets the stage for future generation of transplantable cells with minimal labor, in controlled, automated and reproducible bioreactor large-scale culture systems. Such systems will be useful as hESC technology moves toward industrial and clinical applications.

In the course of our studies on the differentiation of hESC clusters in suspension into neural spheres[14], we observed that in Neurobasal medium hESCs remained undifferentiated in a considerable number of the floating clusters. We therefore sought to develop the Neurobasal medium into a suspension culture system for hESCs. We supplemented Neurobasal medium with Knockout serum replacement (KO-SR), commonly used in hESC culture media, and with Nutridoma-CS, a serum replacement designed to promote the proliferation of cells in suspension[15]. The unique component in Nutridoma-CS is beta-D xylopyranose, a ring-shaped sugar originating from xylose. Supplementation of serum replacement with this sugar allows the culture of cells in suspension without serum[16]. In addition, we supplemented the medium with ECM components previously shown to support the cultivation of undifferentiated hESCs in feeder-free culture systems[3,17,18] and with neurotrophic factors (NT-3, NT-4 and brain-derived neurotrophic factor (BDNF)), which were reported to promote hESC survival[19]. Finally, we included fibroblast growth factor (FGF2) and activin A, as accumulating evidence suggests that FGF and activin signaling play a central role in maintaining the pluripotent state of hESCs[18,20,21] (see Online Methods for complete medium composition).

We evaluated the ability of the key components in the culture system to promote proliferation of undifferentiated hESC clusters in suspension. HESC colonies were dissociated from the feeders and cultured for 3 weeks as floating aggregates in either KO medium or Neurobasal medium, each of which was supplemented with a combination of KO-SR, ECM components, FGF2 and activin A. The KO medium poorly supported the proliferation of hESCs, and the cells tended to differentiate (**Supplementary Fig. 1a**,**b**). In contrast, the Neurobasal medium was more effective for expansion of undifferentiated hESCs. Further supplementation of the Neurobasal medium with Nutridoma-CS significantly ($P < 0.01$) increased the total number of cells obtained after 3 weeks of suspension culture without affecting the percentage of undifferentiated cells (**Supplementary Fig. 1a**,**b**).

Analysis of the effect of ECM components in the culture system showed that they significantly ($P < 0.005$) increased the number of cells without affecting the level of differentiation (**Supplementary Fig. 1c**). Fibronectin and laminin each had a significant ($P < 0.05$) effect on the number of cells, as did the combination of laminin, fibronectin and gelatin ($P < 0.005$). Immunostaining showed increased immunoreactivity with anti-laminin and anti-fibronectin antibodies of clusters that were cultured with the corresponding ECM components

[1]The Hadassah Human Embryonic Stem Cell Research Center, The Goldyne Savad Institute of Gene Therapy, [2]Department of Obstetrics and Gynecology and [3]Department of Genetics, Hadassah-Hebrew University Medical Center, Jerusalem, Israel. Correspondence should be addressed to B.R. (benjaminr@ekmd.huji.ac.il).

Received 24 July 2009; accepted 16 February 2010; published online 28 March 2010; doi:10.1038/nbt.1616

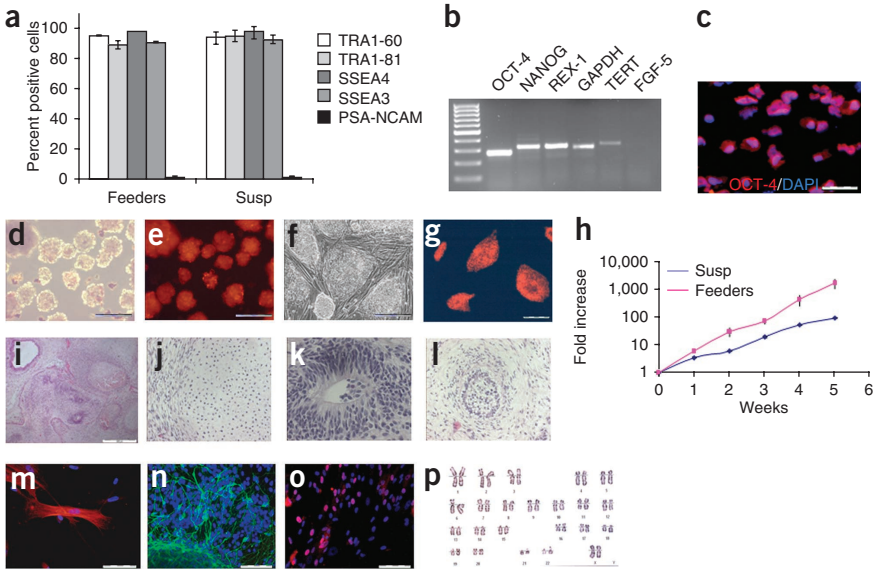


Figure 1 Human ESCs remain pluripotent after 10 weeks propagation in suspension. (**a**) FACS analysis of HES1 cells after cultivation in suspension and on feeders showing that in both culture conditions >90% of the cells express markers of pluripotent stem cells, whereas <2% express PSA-NCAM, which is a marker of early neural differentiation ($n = 3$). Data are presented as mean ± s.d. (**b**) RT-PCR analysis of free-floating clusters of hESCs confirming the expression of transcripts of markers of pluripotency, whereas the expression of the primitive ectoderm marker FGF5 is not detected. (**c**) Immunostaining of cells dissociated from the clusters and plated for 24 h, demonstrating that the majority of cells express OCT-4. (**d**) Darkfield micrograph of the clusters of hESCs in suspension. (**e**) Alkaline phosphatase activity within the hESC aggregates is demonstrated (fluorescence image). (**f**,**g**) After plating of the clusters on feeders, they give rise to colonies with morphological characteristics of colonies of undifferentiated hESCs (**f**, phase contrast image), which are comprised of cells harboring alkaline phosphatase activity (**g**, fluorescence image). (**h**) Proliferation curve showing the fold increase in the number of cells cultivated in suspension or as colonies on feeders. Cell number was monitored each week and the cumulative fold increase in cell number from the starting population was calculated in three independent experiments. (**i**–**l**) Histological analysis of H&E-stained sections of teratoma tumors that developed after inoculation of the hESC-clusters under the testes capsule of NOD/SCID mice showing differentiated tissues representing the three germ layers (low magnitude image (**i**), cartilage (**j**), neural rosette (**k**) and glandular structure (**l**)). (**m**–**o**) Immunostaining of *in vitro*–differentiated progeny, representing the three embryonic germ layers, within the outgrowth of plated embryoid bodies (human muscle actin (**m**), and SOX-17, (**o**)) and neural spheres (β-III Tubulin, (**n**)). (**p**) G-banding analysis showing a normal karyotype. Nuclei are counterstained by DAPI in (**c**) and (**m**–**o**). Scale bars, 20 μm (**c**,**k**,**m**–**o**); 50 μm (**d**,**e**,**j**,**l**); 100 μm (**f**,**g**); 500 μm (**i**).

(**Supplementary Fig. 1d**), suggesting binding of the dissolved ECM components by the hESC clusters. We used the combination of all three ECM components in all subsequent experiments, although it is possible that a less complex ECM composition (e.g., laminin only) would have been sufficient. We also continued to supplement with neurotrophic factors, although they were probably dispensable (**Supplementary Fig. 1e**,**f**).

Next, we used the suspension culture conditions to propagate three hESC lines (HES1, HES2 and H7 (refs. 1,2)) for 10 weeks. The clusters of hESCs formed spheroid aggregates that gradually grew in size and were weekly triturated into small clusters during passaging. During the last 3 weeks of propagation, the hESCs were characterized (**Fig. 1** and **Supplementary Figs. 2** and **3**). Fluorescence-activated cell sorting (FACS) analysis showed that ≥90% of the cells in all three lines expressed markers of pluripotent stem cells, including SSEA-4, TRA-1-60, TRA-1-81 (**Fig. 1a** and **Supplementary Figs. 2a**, **3a** and **4**) and SSEA-3 (**Fig. 1a**; HES1). Immunostaining showed that most of the cells expressed OCT-4 (**Fig. 1c** and **Supplementary Fig. 2b**). The floating clusters expressed alkaline phosphatase (**Fig. 1e**). The percentage of differentiating cells expressing PSA-NCAM (a marker of early neural differentiation) was ≤2% in all three lines (**Fig. 1a** and **Supplementary Figs. 2a**, **3a** and **4**). RT-PCR analysis showed the expression of pluripotent stem cell markers, including OCT-4, NANOG, REX-1 and TERT, lack of expression of FGF5, a marker of primitive ectoderm (**Fig. 1b**), and lack of or low expression of markers of endoderm, mesoderm and extra-embryonic endoderm (**Supplementary Fig. 5a**). Real-time PCR analysis confirmed the high expression levels of OCT-4, NANOG and REX-1 in hESCs cultured as floating clusters similar to monolayer colonies on feeders. It also confirmed the low levels of expression of markers of differentiation (**Supplementary Fig. 5b**). When replated on human feeders, the cell clusters gave rise to colonies comprised of cells with the morphology of undifferentiated hESCs and that expressed alkaline phosphatase (**Fig. 1f**,**g**).

We next demonstrated that hESCs cultivated in suspension maintain their pluripotent potential. Clusters of hESCs from the three lines that were inoculated under the testicular capsule of nonobese diabetic/severe combined immunodeficient (NOD/SCID) mice gave rise to teratomas. Histological analysis of the tumors demonstrated differentiated progeny of the three germ layers (**Fig. 1i–l** and **Supplementary Figs. 2c–f** and **3b–e**). To confirm the pluripotent potential *in vitro,* we cultured the clusters from the three lines under conditions that promote neural spheres[14] or embryoid body[22] formation. After plating, immunostaining showed differentiated progeny representing the three germ layers (**Fig. 1m–o** and **Supplementary Figs. 2g–i** and **3f–h**).

The karyotypes of the three lines were normal after 8 weeks in suspension (**Fig. 1p** and **Supplementary Figs. 2j** and **3i**).

We examined the potential of the suspension culture system to promote undifferentiated cultivation (H7 and HES1 lines) for prolonged periods (20 weeks). FACS analysis showed that the percentage of cells expressing markers of pluripotent stem cells remained ≥90%, and the level of background differentiation was low (≤1% PSA-NCAM$^+$ cells; **Supplementary Fig. 6a**). The karyotype of H7 was normal, whereas an abnormal karyotype was identified in HES1 (**Supplementary Fig. 6b**). The potential of hESCs to develop karyotypic abnormalities (in chromosome 1 of HES1 (ref. 23)) in a variety of culture systems has been described[24,25]. It is unclear whether these abnormalities are related to suboptimal culture conditions, the method of passaging, or a generic property of any type of cell that is propagated for prolonged periods with varying susceptibility to karyotypic changes in different cell lines. It is possible that with further development of the culture system, genetic stability could be improved.

The suspension culture system supported the expansion of the hESCs, though at a considerably lower rate compared with feeder-dependent monolayer hESC cultures. (**Fig. 1h**). We therefore characterized the proliferation, doubling time and cell death in the suspension culture system. FACS

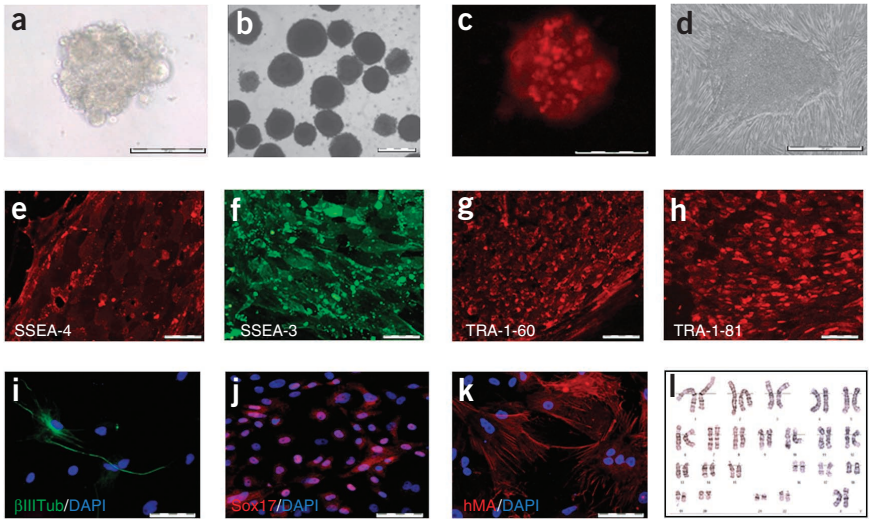


Figure 2 Derivation of hESCs in suspension. (**a**,**b**) Darkfield micrograph of an inner cell mass after transfer to suspension culture conditions (**a**), and of the clusters of cells that were derived from the inner cell mass after 10 weeks of cultivation (**b**). (**c**) Fluorescence image showing alkaline phosphatase activity within a cluster. (**d**–**h**) After plating on feeders, the clusters gave rise to colonies with morphological characteristics of colonies of undifferentiated hESCs (**d**, phase contrast image), which were comprised of cells immunoreactive with anti-SSEA-4 (**e**), SSEA-3 (**f**), TRA-1-60 (**g**) and TRA-1-81 (**h**) (fluorescence images). (**i**–**k**) Immunostaining of *in vitro*–differentiated progeny, representing the three embryonic germ layers, within the outgrowth of plated embryoid bodies (β-III tubulin, (**i**); SOX-17, (**j**); human muscle actin, (**k**)). (**l**) G-banding analysis showing a normal karyotype after 10 weeks of cultivation in suspension. Nuclei are counterstained by DAPI in **i**–**k**. Scale bars, 20 μm (**a**, **e**–**k**); 50 μm (**c**); 100 μm (**b**,**d**). HAD17 hESC line.

analysis of BrdU incorporation showed that 47.8% ± 6.3 (mean ± s.d.) of the cells in suspension were $BrdU^+$ compared to 40.18% ± 3.1 of cells in monolayer cultures ($n = 3$). The doubling time, determined by analyzing the decay of carboxyfluorescein diacetate succinimidyl ester (CFSE) labeling[26] was ~24 h (**Supplementary Fig. 7b**), similar to the doubling time of these hESCs cultured on feeders (23.5 h). The percentage of annexin-V^+ apoptotic cells was 4.6% ± 2.9 in suspension compared to 5.8% ± 2 among hESCs cultivated on feeders, and the percentage of propidium iodide (PI)-positive necrotic cells was 11.3% ± 2.4 and 10.1% ± 5.9, respectively (**Supplementary Fig. 7c**). Thus, the levels of proliferation, cell death and apoptosis in suspension cultures were similar to those in standard feeder-dependent culture conditions.

Given the similar doubling time and cell death rate in our suspension and monolayer cultures, we explored the contribution of cell loss during passaging to the lower expansion rate in suspension. Passaging of the suspension cultures was performed weekly by mechanical trituration into smaller aggregates and further culture for 48 h in the presence of Rho-associated kinase (ROCK) inhibitor[27] to reduce cell death. After mechanical passaging, cell loss was 58.3% ± 5.6 with the suspension culture system, compared to 21.4 ± 7.7 with monolayer cultures ($P < 0.001$). Hence, cell loss during passaging was the major cause of the lower expansion rate in suspension. This problem might be alleviated by refinement of the passaging method.

We evaluated whether the suspension culture conditions can support the derivation of new hESC lines. Sixteen human blastocysts, diagnosed as carriers of genetic abnormalities and destined to be discarded, were obtained. The inner cell masses were isolated with the assistance of a laser[22] from 15 blastocysts. The inner cell masses and an additional intact blastocyst were cultured in the suspension culture conditions. The cells of two of the inner cell masses from embryos with neurofibromatosis, and of the intact embryo, affected by spastic paraplegia 4, proliferated. Each of these three cell clusters were mechanically dissociated into multiple cell aggregates. The three newly derived cell lines (HAD17–19) were further expanded for an additional 10 weeks (**Fig. 2** and **Supplementary Figs. 8** and **9**). The morphological characteristics of the clusters of the three cell lines were similar to those of clusters of established cell lines. The cells within the newly derived clusters expressed alkaline phosphatase and markers of pluripotent

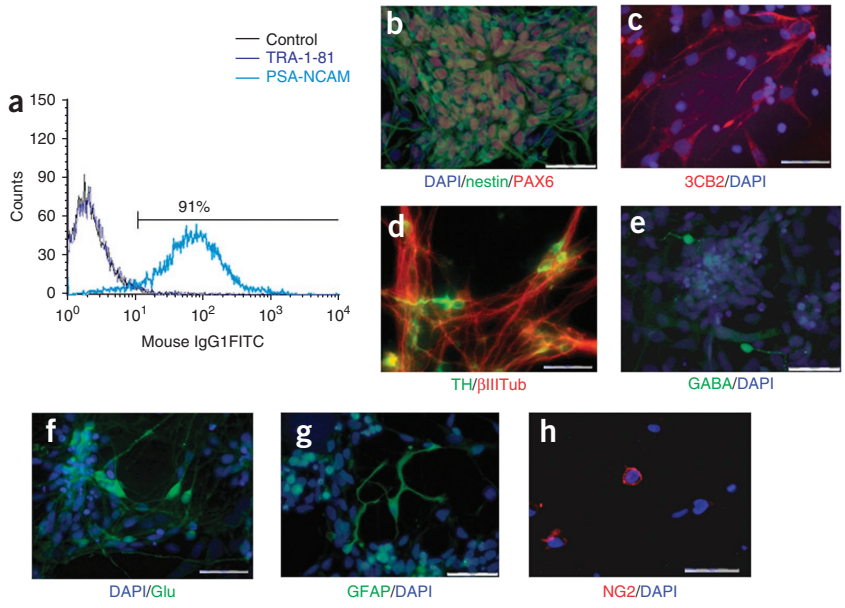


Figure 3 Controlled conversion of the hESC clusters in suspension into neural precursor spheres. Clusters of H7 cells, cultivated in suspension for 7 weeks, were transferred and further cultured 4 weeks in a chemically defined medium supplemented with noggin and FGF2. (**a**) FACS analysis of one representative experiment showing that 91% of the cells expressed PSA-NCAM, whereas 1.2% expressed TRA-1-81. (**b**–**h**) After plating and culturing for 1 week on laminin, indirect immunofluorescence staining showed cells within rosettes expressing markers of neural precursors such as nestin and Pax6 (**b**), the neural stem/radial glial cell marker 3CB2 (**c**), subtypes of neurons expressing β-III tubulin and tyrosine hydroxylase (TH) (**d**), GABA (**e**) and glutamate (**f**), as well as cells expressing the astrocyte marker GFAP (**g**) and the marker of oligodendroglial progenitors NG2 (**h**). Nuclei are counterstained by DAPI. Scale bars, 20 μm.

stem cells. Upon plating on human feeders, they gave rise to colonies and cells with morphological characteristics of undifferentiated hESCs, which expressed markers of pluripotent stem cells, including TRA-1-60, TRA-1-81, SSEA-3 and SSEA-4. Upon differentiation of the free-floating clusters *in vitro* into embryoid bodies and plating, all three lines differentiated into progeny representing the three germ layers. The karyotype at passage 10 was normal in all lines. The clusters were frozen and successfully thawed by vitrification[22]. Although the suspension culture system could promote the derivation of hESC lines, it did not allow clonal expansion. The successful derivation of new hESC lines in suspension demonstrated the robustness of the culture system in promoting proliferation of undifferentiated hESCs. These results showed that feeders or attachment to an ECM layer are not required to develop hESC lines and that new lines can be derived in suspension conditions.

Lastly, we determined whether the hESC free-floating clusters could be directed to differentiate in suspension into clusters enriched for cells of a specific lineage. For this purpose, clusters of HES1 and H7 hESC lines (after 10 weeks cultivation) were cultured for 4 weeks in a chemically defined medium supplemented with FGF2 and noggin, as described[14]. The clusters gradually acquired the morphology of neural spheres, and FACS analysis showed that ≥90% of the cells expressed the neural marker PSA-NCAM, whereas the expression of TRA-1-81 was downregulated (**Fig. 3a**). We next demonstrated the potential of the neural precursors within the spheres to differentiate into progeny representing the three neural lineages. After plating the neural precursors on laminin and culturing for a week in the presence of survival factors, they differentiated into neurons expressing β-III tubulin, tyrosine hydroxylase (TH), GABA and glutamate (**Fig. 3b–h** and **Supplementary Fig. 10**). The neural precursors also gave rise to glial fibrillary acidic protein (GFAP)-expressing astrocytes and NG2-expressing oligodendrocyte progenitors (**Fig. 3g,j**). After 1 week of differentiation, the cultures also included remnants of neural precursors in rosettes co-expressing PAX-6 and nestin, and neural stem/radial glial cells expressing 3CB2 (**Fig. 3b,c** and **Supplementary Fig. 10a,b**). When the plated spheres were differentiated in the presence of FGF8 and sonic hedgehog (SHH) (which promote midbrain dopamine neuronal differentiation[28]), multiple neurons co-expressing engrailed-1 and TH were observed (**Supplementary Fig. 10f**). Co-expression of these markers is characteristic of midbrain dopaminergic neurons[29]. Thus, the clusters of hESCs could be directed to differentiate in suspension into a population highly enriched for precursors of a specific lineage.

In conclusion, we have demonstrated that pluripotent hESCs can be derived, propagated and directed to differentiate into neural precursors in a feeder-free suspension culture system of floating clusters. This is an important step toward developing controlled suspension systems for large-scale expansion followed by directed differentiation in bulk of hESCs.

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 are grateful to the following members of the Hadassah Human Embryonic Stem Cell Research Center: M. Gropp, M. Aharonowiz and O. Singer for technical assistance; S. Tennenbaum for editing the manuscript. We thank K.M. Yamada (National Institute of Dental and Craniofacial Research, National Institutes of Health) for providing anti-fibronectin antibody, N. Benvenisty for the QPCR primers and WiCell Research Institute for providing H7 hESCs. This research was supported by a gift from Judy and Sidney Swartz, the Sidney Swartz Chair in Human Embryonic Stem Cell Research and Legacy Heritage Fund.

AUTHOR CONTRIBUTIONS

D.S. designed and performed the experiments, analyzed the data and wrote the manuscript; H.K. and M.C. performed the neural differentiation study; S.E.-R. conducted immunostainings and confocal analysis; Y.G. performed the teratoma studies; P.I. contributed to developing the concept of suspension culture; T.T. contributed to the experiments; M.I. performed PCR analysis; E.A. contributed to embryo recruitment, culture and isolation of inner cell masses; R.R. and Y.B.-Z. conducted karyotype analysis. B.R. conceived the study and wrote the paper.

COMPETING FINANCIAL INTERESTS

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. Reubinoff, B.E., Pera, M.F., Fong, C.Y., Trounson, A. & Bongso, A. Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. *Nat. Biotechnol.* **18**, 399–404 (2000).
2. Thomson, J.A. *et al.* Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145–1147 (1998).
3. Ludwig, T.E. *et al.* Derivation of human embryonic stem cells in defined conditions. *Nat. Biotechnol.* **24**, 185–187 (2006).
4. Richards, M., Fong, C.Y., Chan, W.K., Wong, P.C. & Bongso, A. Human feeders support prolonged undifferentiated growth of human inner cell masses and embryonic stem cells. *Nat. Biotechnol.* **20**, 933–936 (2002).
5. Xu, C. *et al.* Feeder-free growth of undifferentiated human embryonic stem cells. *Nat. Biotechnol.* **19**, 971–974 (2001).
6. Braam, S.R. *et al.* Feeder-free culture of human embryonic stem cells in conditioned medium for efficient genetic modification. *Nat. Protoc.* **3**, 1435–1443 (2008).
7. Wang, L. *et al.* Self-renewal of human embryonic stem cells requires insulin-like growth factor-1 receptor and ERBB2 receptor signaling. *Blood* **110**, 4111–4119 (2007).
8. Kurosawa, H. Methods for inducing embryoid body formation: *in vitro* differentiation system of embryonic stem cells. *J. Biosci. Bioeng.* **103**, 389–398 (2007).
9. McDevitt, T.C. & Palecek, S.P. Innovation in the culture and derivation of pluripotent human stem cells. *Curr. Opin. Biotechnol.* **19**, 527–533 (2008).
10. Lock, L.T. & Tzanakakis, E.S. Expansion and differentiation of human embryonic stem cells to endoderm progeny in a microcarrier stirred-suspension culture. *Tissue Eng. Part A* **15**, 2051–2063 (2009).
11. Nie, Y., Bergendahl, V., Hei, D.J., Jones, J.M. & Palecek, S.P. Scalable culture and cryopreservation of human embryonic stem cells on microcarriers. *Biotechnol. Prog.* **25**, 20–31 (2009).
12. Oh, S.K. *et al.* Long-term microcarrier suspension cultures of human embryonic stem cells. *Stem Cell Res. (Amst.)* **4**, 4 (2009).
13. Krawetz, R. *et al.* Large-scale expansion of pluripotent human embryonic stem cells in stirred suspension bioreactors. *Tissue Eng. Part C Methods* **8**, 8 (2009).
14. Itsykson, P. *et al.* Derivation of neural precursors from human embryonic stem cells in the presence of noggin. *Mol. Cell. Neurosci.* **30**, 24–36 (2005).
15. Hoover, C.S. & Martin, R.L. Antibody production and growth of mouse hybridoma cells in Nutridoma media supplements. *Biotechniques* **8**, 76–82 (1990).
16. Stockinger, H. Serum-free medium for mammalian cells. US patent 5,063,157 (1991).
17. Amit, M., Shariki, C., Margulets, V. & Itskovitz-Eldor, J. Feeder layer- and serum-free culture of human embryonic stem cells. *Biol. Reprod.* **70**, 837–845 (2004).
18. Xu, C. *et al.* Basic fibroblast growth factor supports undifferentiated human embryonic stem cell growth without conditioned medium. *Stem Cells* **23**, 315–323 (2005).
19. Pyle, A.D., Lock, L.F. & Donovan, P.J. Neurotrophins mediate human embryonic stem cell survival. *Nat. Biotechnol.* **24**, 344–350 (2006).
20. Beattie, G.M. *et al.* Activin A maintains pluripotency of human embryonic stem cells in the absence of feeder layers. *Stem Cells* **23**, 489–495 (2005).
21. Furue, M.K. *et al.* Heparin promotes the growth of human embryonic stem cells in a defined serum-free medium. *Proc. Natl. Acad. Sci. USA* **105**, 13409–13414 (2008).
22. Turetsky, T. *et al.* Laser-assisted derivation of human embryonic stem cell lines from IVF embryos after preimplantation genetic diagnosis. *Hum. Reprod.* **23**, 46–53 (2008).
23. Herszfeld, D. *et al.* CD30 is a survival factor and a biomarker for transformed human pluripotent stem cells. *Nat. Biotechnol.* **24**, 351–357 (2006).
24. Baker, D.E. *et al.* Adaptation to culture of human embryonic stem cells and oncogenesis *in vivo*. *Nat. Biotechnol.* **25**, 207–215 (2007).
25. Mitalipova, M.M. *et al.* Preserving the genetic integrity of human embryonic stem cells. *Nat. Biotechnol.* **23**, 19–20 (2005).
26. Lyons, A.B. & Parish, C.R. Determination of lymphocyte division by flow cytometry. *J. Immunol. Methods* **171**, 131–137 (1994).
27. Watanabe, K. *et al.* A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nat. Biotechnol.* **25**, 681–686 (2007).
28. Lee, S.H., Lumelsky, N., Studer, L., Auerbach, J.M. & McKay, R.D. Efficient generation of midbrain and hindbrain neurons from mouse embryonic stem cells. *Nat. Biotechnol.* **18**, 675–679 (2000).
29. Yan, Y. *et al.* Directed differentiation of dopaminergic neuronal subtypes from human embryonic stem cells. *Stem Cells* **23**, 781–790 (2005).

ONLINE METHODS

HESC feeder-dependent culture system. Human ESCs of HES1, HES2 (ref. 1) and H7 (ref. 2) lines at passages 24–33, 24–27 and 38–45, respectively, with normal karyotypes, were cultured on foreskin human feeders in KO medium consisting of 85% KO-DMEM, 15% KO-SR, 1 mM glutamine, 0.1 mM β-mercaptoethanol, 1% nonessential amino acids, 50 units/ml penicillin, 50 μg/ml streptomycin (Invitrogen) and 4 ng/ml FGF2 (Peprotech), as previously described[30]. HES1 and HES2 cells were passaged weekly with Ca/Mg^{2+}-free PBS supplemented with 0.05% EDTA (Biological Industries) or type IV collagenase (1 mg/ml; Invitrogen), whereas mechanical dissociation was used for passaging H7. To determine expansion rate, in each passage, colonies from one well were mechanically removed, dissociated by trituration, and replated on feeders in four wells. To determine average cell number per well, 7 d after passage, the colonies in each of three wells were disaggregated and counted and the average number of cells per well was calculated. The fourth well was further passaged as above. To determine cell loss during passaging, half of the colonies that were removed from each well were triturated as during cell passaging. Cell number was compared between the triturated and nontriturated clusters 2 h after plating.

Suspension culture system for propagating hESC. hESC colonies were dissociated with collagenase IV (1 mg/ml, 1.5–2 h at 37 °C) and by agitation of the culture plate. The colonies were triturated into small cell clusters, which were suspended within nonadherent 12-well tissue culture dishes (Costar, Corning) at a density of ~$0.7–1.2 \times 10^6$ cells/ml in Neurobasal medium, 14% KO-SR, L-glutamine 2 mM, 50 units/ml penicillin, 50 μg/ml streptomycin, 1% nonessential amino acids (all from Invitrogen). The medium was supplemented with FGF2 20 ng/ml, activin A 25 ng/ml (both from Peprotech), fibronectin 1 μg/ml (BD Biosciences), laminin 0.5 μg/ml, gelatin 0.001% (both from Sigma), the neurotrophins BDNF, NT3 and NT4, 10 ng/ml each (Peprotech) and 1× Nutridoma-CS (Roche). Aggregation and overgrowth of clusters was occasionally prevented by trituration with a 1,000 μl pipettor tip as required. Once a week, the clusters from each well were partially disaggregated by gentle trituration and split into two wells for further culture in fresh medium supplemented with ROCK inhibitor (Sigma)[27] 10 μM, which was removed after 48 h. Trituration could be replaced with similar results by using stem cell passaging tool (Invitrogen). The medium was changed every other day. For medium replacement, every other day, the tissue culture dishes were tilted under binocular microscope until cell clusters settled at the lowest part of each well and 80% of the medium was gently replaced. To determine average cell number per well, 7 d after passage the clusters in each of three wells were disaggregated and counted and the average number of cells per well was calculated. To determine cell loss during passaging, half of the clusters in each well were triturated as during cell passaging. Cell number was compared between the triturated and nontriturated clusters after 2 h.

FACS and alkaline phosphatase activity analysis. The hESC clusters were dissociated with PBS solution containing 2.25 mM EDTA and 0.06% trypsin, for 10 min at 37 °C, followed by gentle trituration to a single-cell suspension. The hESCs were then washed with PBS supplemented with 1% BSA and 0.05% sodium azide (both from Sigma). The cells were incubated with anti-SSEA4 (1:100, mouse monoclonal IgG3; Developmental Studies Hybridoma Bank (DHSB)), anti-TRA-1-60 (1:100, monoclonal mouse IgM; Chemicon), anti-TRA-1-81 (1:100, monoclonal mouse IgM; Chemicon), anti-SSEA3 (1:100, monoclonal rat IgM; Chemicon), and anti-PSA-NCAM (1:100, monoclonal mouse IgM; Chemicon). Control hESCs were stained with respective isotype control antibodies. Primary antibodies were detected using fluorescein isothiocyanate (FITC)-labeled goat anti-mouse immunoglobulin (1:100, Dako) or Alexa Fluor-labeled goat anti-rat IgM (1:100, Invitrogen). PI was added (final concentration of 4 μg/ml) for better gating of viable cells. FACS analysis was performed using the FACSCalibur system (Becton, Dickinson). Alkaline Phosphatase Substrate kit I (Vector) was used to analyze alkaline phosphatase activity within clusters according to the manufacturer's instructions.

Replating of hESCs cultivated in suspension on feeders. Floating aggregates of hESCs were triturated with a 1,000 μl pipettor tip into small clusters that were plated on fresh feeders, cultured in KO-SR medium supplemented with 4 ng/ml FGF2, and the resulting colonies were passaged routinely as described above. Alkaline phosphatase activity within colonies was analyzed as for clusters.

Differentiation *in vitro*. For embryoid body formation, clusters of undifferentiated hESCs were transferred and cultured for 3–4 weeks in DMEM medium (Invitrogen), supplemented with 20% FBS (Biological Industries), 1 mM L-glutamine, 0.1 mM β-mercaptoethanol, 1% nonessential amino acid stock, 50 units/ml penicillin, 50 μg/ml streptomycin. For further differentiation, the embryoid bodies were dissociated using trypsin (0.025%, 3 mM EDTA in PBS) and plated on poly-D-lysine (30–70 kDa, 10 μg/ml; Sigma) and laminin (4 μg/ml; Sigma) precoated glass coverslips for an additional 1–2 weeks culture in the same medium.

For controlled differentiation into neural spheres, the hESC clusters were transferred and cultured for 3 weeks in DMEM/F12 medium (Invitrogen) supplemented with B27 (1%, Invitrogen), FGF2 (20 ng/ml) and noggin (600 ng/ml, R&D Systems), followed by 1 week in the presence of FGF2 without noggin.

For further differentiation, the neural spheres were triturated to small clusters and plated on poly-D-lysine and laminin-coated glass coverslips and cultured for an additional week with DMEM/F12/B27 medium in the presence of survival factors (ascorbic acid (200 μM, Sigma), NT-4 and BDNF (10 ng/ml each, Peprotech)). For midbrain differentiation, the plated partially disaggregated spheres were cultured in the presence of 100 ng/ml of FGF8 and 200 ng/ml SHH (both from R&D Systems) for a week followed by further differentiation for a week in the presence of survival factors as above.

Immunostaining. Cells were fixed with 4% paraformaldehyde for 20 min at 23 °C. For immunostaining with anti-intracellular marker antibodies, cell membranes were permeabilized with 0.2% Triton X100 (Sigma) in PBS for 5 min. The cells were incubated with the following primary antibodies: anti-OCT-3/4 (mouse IgG, 1:100, Santa Cruz Biotechnology), anti-β-III-tubulin (mouse monoclonal IgG2b, 1:2,000, Sigma), anti-TH (1:200, Pel Freeze), anti-GABA (1:1,000, Sigma), anti-glutamate (1:2,000, Sigma), anti-3CB2 (1:100, DHSB), anti-GFAP (1:100, Dako), anti-NG2 (1:50, Chemicon), anti-human muscle actin (1:50, Dako), anti-SOX-17 (1:50, R&D Systems). Primary antibody localization was performed by using FITC-conjugated goat anti-rabbit immunoglobulins (1:20-50, Dako), or goat anti-mouse immunoglobulin conjugated with Cy3 (1:500, Jackson ImmunoResearch). Mounting medium containing 4′,6-diamidino-2-phenylindole (DAPI; Vector) was used for nuclei counterstaining and the specimens were visualized with a Nikon E600 fluorescent microscope.

For immunostaining of ECM components, clusters of hESCs cultivated in the presence or absence of ECM components were permeabilized in 0.5% Triton X-100/4% paraformaldehyde and fixed in a 4% paraformaldehyde/PBS-5% sucrose solution. The clusters were stained in suspension with anti-fibronectin (rabbit polyclonal antibody, a gift from K.M. Yamada) and anti-human laminin (rat monoclonal antibody, 1:100, Chemicon). Primary antibodies were detected using RhodamineX-conjugated donkey anti-Rabbit IgG and Cy2-conjugated donkey anti-Mouse IgG (Jackson Immunoresearch Labs). Imaging performed by quadruple laser-assisted confocal microscopy (FluoView FV100, Olympus), with a 40× objective UPLAPO40xOI2.

Teratoma formation. The potential of the hESCs to form teratoma tumors was evaluated in NOD/SCID mice. These experiments were approved by the Institutional Committee for animal research of the Hebrew University/Hadassah Medical School. Cluster of hESCs were injected under the testicular capsule of 6-week-old NOD/SCID mice (Harlan) (30–40 clumps per testis). Eight to twelve weeks later, the resulting tumors were removed, embedded in paraffin and sections were stained with H&E. The specimens were visualized with a Nikon TE300 microscope with a 20× objective.

Derivation of hESCs in suspension. *In vitro*–fertilized human embryos that were not diagnosed in the preimplantation period as genetically normal were recruited for the study subject by informed consent of the couples. The study was approved by the ethics committee at the Hadassah Medical Center as well as the Israeli Ministry of Health National Helsinki Committee for Genetic Research in Humans. Fifteen inner cell masses of abnormal blastocysts were isolated by a laser-assisted system[22] and transferred to the suspension culture

conditions as above. The medium was supplemented with ROCK inhibitor 10 μM during the initial 7 d. One whole embryo was also plated in the same conditions. Proliferating clusters were mechanically dissected into smaller aggregates for further propagation in suspension.

Analysis of apoptosis necrosis and proliferation. Proliferation was determined by FACS analysis of BrdU incorporation using the BrdU Flow Kit (BD Biosciences) after 4 h incubation with BrdU according to the manufacturer's instructions. Cell division was monitored according to the technique described by Lyons and Parish (1994) with slight modifications. Clusters of hESCs were incubated in PBS supplemented with CFSE at a final concentration of 5 μM at 23 °C for 15 min. CFSE-labeled cells were washed three times with Neurobasal medium and cultured for up to 72 h. At 24 h intervals, samples of clusters were harvested for FACS analysis (as above) of CFSE fluorescence intensity. Cell apoptosis was measured with the Phosphatidyl Serine Detection kit (IQ Products) according to the manufacturer's instructions. Cells were washed in calcium buffer and incubated with FITC-conjugated anti-annexin V antibodies for 20 min. Propidium iodide (PI) was added to label nonviable cells and FACS analysis was performed as above.

Karyotype. For karyotype analysis, hESC clusters were reseeded on feeders and expanded with collagenase type IV for 3–4 passages. The colonies were incubated for 40 min with 0.2–0.3 μg/ml demecolcine (Sigma). The hESC colonies were then removed from the feeders, dissociated with 0.05% EDTA, centrifuged at 170*g* for 5 min, resuspended in 0.075 M KCl (Sigma) and incubated for 10 min in 37 °C, followed by fixation with 3:1 methanol/acetic acid. The karyotype of 10–20 metaphases was analyzed using the G-banding method.

RT-PCR. Total RNA was isolated using TRI-Reagent (Sigma). cDNA synthesis was carried out using Moloney murine leukemia virus reverse transcriptase (M-MLV RT) and random primers, according to the manufacturer's instructions (Promega). PCR was carried out using Taq DNA Polymerase (Gibco-BRL) with denaturation, 94 °C; 30 s, annealing, 55 °C; 30 s, extension, 72 °C; 45 s for 40 cycles. Primer sequences (forward and reverse 5′–3′) and the length of the amplified products were as follows:

hOCT4 (AGTGAGAGGCAACCTGGAGA, GTGAAGTGAGGGCTCCCATA; 273 bp);
hNANOG (CGAAGAATAGCAATGGTGTGACG, TTCCAAAGCAGCCTCCA AGTC ; 328 bp);
hREX-1 (GAGCCTGTGTGAACAGAAC, CATAGCACACATAGCCATC; 322 bp);
hTERT (CTGCAGCTCCCATTTCAT, GGATGGTCTTGAAGTCTG; 306 bp);
hFGF5 (GATCCCACGAAGCCAATA, GCTCCGACTGCTTGAATC; 338 bp);
hCG (GTCAACACCACCATCTGTGC; GGCCTTTGAGGAAGAGGAGT; 285 bp);
αFP (CCATGTACATGAGCACTGTTG;CTCCAATAACTCCTGGTATCC; 338 bp);
FOXA2 (ACTGTGTAGACTCCTGCTTCTTC;GCACGCAGAAACCATAAAT; 305 bp);
SOX17 (CGCACGGAATTTGAACAGTA;GGATCAGGGACCTGTCACAC; 180 bp);
Brachyury (CTTCCCTGAGACCCAGTTCA;CAGGGTTGGGTACCTGTCAC; 289 bp);
Goosecoid (TTCCAGGAGACCAAGTACCC,TCGTCTGTCTGTGCAAGTCC; 298 bp);
hGAPDH (GACAACAGCCTCAAGATC, GTCCACCACTGACACGTT; 311 bp)

For Q-PCR TaqMan Assays-on-Demand Gene Expression Products (OCT4, Hs01895061_μ1; NANOG, Hs02387400_g1; REX1, Hs00399279_m1; FGF5, Hs00738132_m1; Brachyury, Hs00610080_m1; Goosecoid, Hs00418279_m1; SOX17, Hs00751752_s1; FOXA2, Hs00232764_m1; hCG, Hs00361224_gH; alphaFP, Hs00173490_m1), TaqMan Universal PCR Master Mix and ABI Prism 7900HT Sequence Detection System (Applied Biosystems) were used. β-glucuronidase (GusB, Hs99999908_m1) was an internal reference for normalization.

Statistical analysis. Data are presented as mean ± s.d. The significance of differences between treatments was calculated using one-tailed *t*-test.

30. Gropp, M. & Reubinoff, B. Lentiviral vector-mediated gene delivery into human embryonic stem cells. *Methods Enzymol.* **420**, 64–81 (2006).

doi:10.1038/nbt.1616

nature biotechnology

Interfamily transfer of a plant pattern-recognition receptor confers broad-spectrum bacterial resistance

Séverine Lacombe[1,5,6], Alejandra Rougon-Cardoso[1,5,6], Emma Sherwood[1,5,6], Nemo Peeters[2], Douglas Dahlbeck[3], H Peter van Esse[4], Matthew Smoker[1], Ghanasyam Rallapalli[1], Bart P H J Thomma[4], Brian Staskawicz[3], Jonathan D G Jones[1] & Cyril Zipfel[1]

Plant diseases cause massive losses in agriculture. Increasing the natural defenses of plants may reduce the impact of phytopathogens on agricultural productivity. Pattern-recognition receptors (PRRs) detect microbes by recognizing conserved pathogen-associated molecular patterns (PAMPs)[1–3]. Although the overall importance of PAMP-triggered immunity for plant defense is established[2,3], it has not been used to confer disease resistance in crops. We report that activity of a PRR is retained after its transfer between two plant families. Expression of EFR (ref. 4), a PRR from the cruciferous plant *Arabidopsis thaliana*, confers responsiveness to bacterial elongation factor Tu in the solanaceous plants *Nicotiana benthamiana* and tomato (*Solanum lycopersicum*), making them more resistant to a range of phytopathogenic bacteria from different genera. Our results in controlled laboratory conditions suggest that heterologous expression of PAMP recognition systems could be used to engineer broad-spectrum disease resistance to important bacterial pathogens, potentially enabling more durable and sustainable resistance in the field.

The ever-growing world population, the threat of climate change and the increasing interest in crop-derived biofuel production are some of the factors that threaten global food security. Microbial diseases and pests place major constraints on food production and agriculture. Agrochemical applications are the most common means of controlling these, but more sustainable methods are required[5]. One way to improve plant disease resistance is to enhance the capability of the plants' own innate immune system[6,7]. Although constitutive physical and chemical barriers contribute to their defense, like all higher eukaryotes, plants also depend for their survival on active recognition of microbial invaders. Plants can recognize potential pathogens via two perception systems[1]. One detects conserved microbial molecules, named pathogen- or microbe-associated molecular patterns (PAMPs or MAMPs), through PRRs, leading to PAMP-triggered immunity (PTI). The other evolved to recognize microbial virulence effectors, usually through intracellular resistance proteins (R proteins), causing effector-triggered immunity (ETI).

ETI corresponds to what is classically referred to as gene-for-gene, vertical or race-specific resistance[1]. It generally occurs between cultivars of a given plant species bearing a particular *R* gene and a limited number of pathogenic strains carrying the matching virulence effector. *R* gene–mediated resistance is widely used in breeding programs to control plant diseases. However, this type of resistance rarely confers broad-spectrum disease resistance. Moreover, it is often rapidly overcome by evolving pathogens that lose or mutate the nonessential recognized effector or that produce new effectors to counteract ETI[1,8,9].

By definition, PAMPs are conserved across a wide range of microbes, which may or may not be pathogenic. Because these molecules are essential for viability or lifestyle, microbes are less likely to evade host immunity through mutation or deletion of PAMPs, compared with virulence effectors. PTI contributes to the plant innate immunity that is activated even during a susceptible (compatible) interaction. More notably, however, it probably constitutes an important aspect of non-host resistance, which accounts for why most plants are resistant to the majority of pathogens they encounter[10–17].

Relatively few plant PRRs have been identified. The leucine-rich repeat receptor kinases FLS2, EFR and XA21 recognize the bacterial PAMPs flagellin, elongation factor Tu (EF-Tu) and the type I–secreted sulfated protein 'activator of XA21-mediated immunity' (Ax21), respectively[2,3,18]. The transmembrane LysM domain–containing protein CeBiP binds the fungal PAMP chitin[2,3]. Perception of both chitin and unknown bacterial PAMP(s) require the LysM receptor kinase CERK1[2,3,19]. The receptor-like proteins EIX1 and EIX2 are PRRs for fungal xylanase, and the soluble protein GBP binds branched 1,6-1,3-β-glucans from the oomycete *Phytophthora sojae*[2,3].

Plant mutants in which PAMP recognition is affected are more susceptible to adapted pathogens (reflecting defects in basal resistance) and allow some degree of disease progression by non-adapted pathogens (reflecting defects in non-host resistance). For example, loss of flagellin recognition in *A. thaliana* and *N. benthamiana* enhances

[1]The Sainsbury Laboratory, Norwich Research Park, Norwich, UK. [2]Centre National de la Recherche Scientifique-Institut National de la Recherche Agronomique, Unité Mixte de Recherche, Laboratoire des Interactions Plantes Microorganismes, Castanet-Tolosan, France. [3]Department of Plant and Microbial Biology, University of California, Berkeley, California, USA. [4]Laboratory of Phytopathology, Wageningen University, Wageningen, The Netherlands. [5]Present addresses: Institut National de la Recherche Agronomique, Unité de Recherche, Génétique et Amélioration des Fruits et Légumes, Montfavet, France (S.L.), Laboratorio Nacional de Genómica para la Biodiversidad, CINVESTAV Irapuato, Irapuato, Mexico (A.R.-C.) and Department of Molecular Microbiology, John Innes Centre, Norwich, UK (E.S.). [6]These authors contributed equally to this work. Correspondence should be addressed to C.Z. (cyril.zipfel@tsl.ac.uk).

Received 9 December 2009; accepted 9 February 2010; published online 14 March 2010; doi:10.1038/nbt.1613

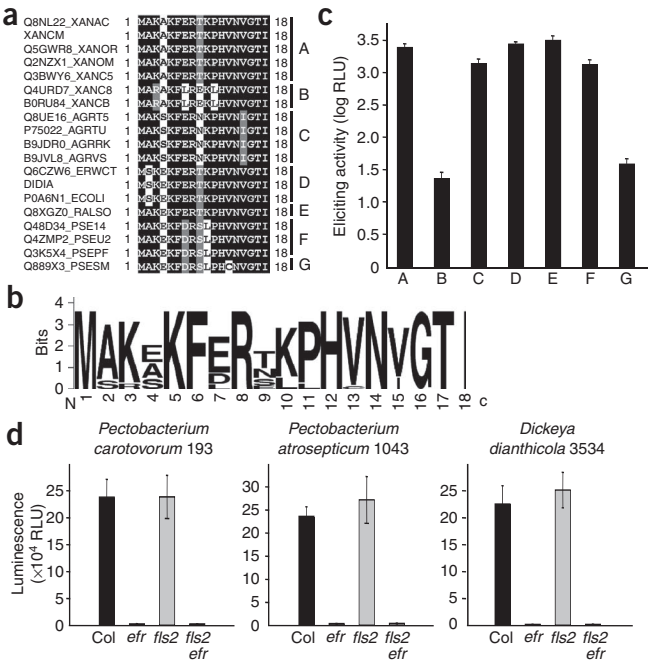


Figure 1 Eliciting activities of elf18 peptides and EF-Tu from selected phytopathogenic bacteria in *A. thaliana*. (**a**) Alignment of elf18 regions from selected bacteria. Capital letters on the right indicate the subgroups of elf18 peptides. Accession numbers are from UniProtKB. XANAC, *X. axonopodis* pv. *citri* 306; XANCM, *X. campestris* pv. *musacearum* 4381; XANOR, *X. oryzae* pv. *oryzae* KXO85; XANOM, *X. oryzae* pv. *oryzae* MAFF 311018; XANC5, *X. campestris* pv. *vesicatoria* 85-10; XANC8, *X. campestris* pv. *campestris* 8004; XANCB, *X. campestris* pv. *campestris* B100; AGRT5, *A. tumefaciens* C58; AGRTU, *A. tumefaciens*; AGRRK, *Agrobacterium radiobacter* K84; AGRVS, *Agrobacterium vitis* S4; ERWCT, *E. carotovora* ssp. *atroseptica*/*P. atrosepticum* SCRI1043; DIDIA, *Dickeya dianthicola*/*E. chrysanthemi* SCRI3534; ECOLI, *E. coli* K12; RALSO, *R. solanacearum* GMI1000; PSE14, *P. syringae* pv. *phaseolicola* 1448A/Race 6; PSEU2, *Pss* B728a; PSEPF, *Pseudomonas fluorescens* Pf0-1; PSESM, *Pto* DC3000. (**b**) WebLogo representation of the elf18 consensus sequence. (**c**) Oxidative burst triggered by elf18 peptides from different subgroups as defined in **a**. We calculated the eliciting activity as the amount of relative light units (RLU) produced in response to 1 μM elf18 peptide minus the amount of ROS produced in response to water in wild-type (Col-0 ecotype) *A.* thaliana leaf discs. Results are averages ± s.e.m. ($n = 12$). (**d**) Oxidative burst triggered by 10 μl bacterial extracts from *P. carotovorum* 193, *P. atrosepticum* 1043 and *D. dianthicola* 3534 in *A. thaliana* leaf discs from wild-type (Col-0; black), *efr* (light gray), *fls2* (gray) and *fls2 efr* (white) plants, measured as RLU. Results are averages ± s.e.m. ($n = 12$). We repeated all experiments at least three times with similar results.

susceptibility to virulent, weakly virulent and non-adapted *Pseudomonas syringae* strains[10–13]. EFR contributes to resistance to *Agrobacterium tumefaciens* and weakly virulent strains of *P. syringae* pathovar *tomato* (*Pto*) DC3000 in *A. thaliana*[4,20]. The rice PRR XA21 confers resistance to *Xanthomonas oryzae* pv. *oryzae*[21].

Although some PAMPs are recognized by many plant species, others have a narrower range of recognition[3]. For example, flagellin is recognized by both monocotyledonous and dicotyledonous species, as reflected by the identification of FLS2 homologs in *A. thaliana*, tomato, *N. benthamiana* and rice[2]. In contrast, EF-Tu (or its eliciting epitope elf18), cold-shock protein (or its eliciting epitope csp22) and Ax21 are only known to be recognized naturally by members of the Brassicaceae, Solanaceae and some rice cultivars, respectively[3,18].

On the basis of this observation, we tested whether the transfer of new PAMP recognition capacities across plant families would confer broad-spectrum disease resistance. We hypothesized that pathogens that are adapted to a given host species might not have evolved virulence effectors targeting PRRs that are normally absent from the host species. We chose the Brassicaceae-specific PRR EFR (ref. 4) because the high level of conservation of EF-Tu protein sequences across bacteria[22] offers the possibility that EFR could confer resistance against a wide range of bacterial pathogens.

To confirm that EFR detects EF-Tu proteins from important phytopathogenic bacteria, we first assessed the variability and eliciting activity of elf18 peptides derived from a selection of phytopathogenic bacteria spanning several genera. The 'classical' elf18 peptide sequence (acetyl-MSKEKFERTKPHVNVGTI) is based on the EF-Tu from *Escherichia coli*[22]. Previous work based on alanine-scanning and deletion analyses has shown that a fully active minimal peptide requires the sequence acetyl-MxKxKFxRxxxxxxxxx (where x is any amino acid)[22]. Alignment of the 19 selected elf18 sequences revealed seven different groups, with a maximum of five residues differing from the *E. coli*–based elf18 (elf18^{Ecoli}; **Fig. 1a**). In agreement with the strong conservation of the consensus elf18 peptide (**Fig. 1b**), synthetic peptides from the seven groups all induced the production of reactive-oxygen species (ROS) in *A. thaliana* wild-type leaves (**Fig. 1c**). However, synthetic peptides from groups B and G (corresponding to strains of *Xanthomonas campestris* pv. *campestris* and *Pto* DC3000, respectively) showed only 0.8%–3.2% of the activity measured with peptides from the other groups (**Fig. 1c**). In the case of group B, this is probably due to the Lys→Arg substitution at position 3 of the elf18 peptide (**Fig. 1a**), a residue that is required for full eliciting activity[22]. Although elf18$^{Pto\text{DC}3000}$ (group G) does not show any variability in key residues, it has five substitutions compared to elf18^{Ecoli} (group D; **Fig. 1a**,**b**), which probably explain its reduced activity in *A. thaliana* leaves (**Fig. 1c**)[22].

Sequence information necessary to predict elf18 peptides may not always be available for all bacterial strains, especially in the case of emerging pathogens. Nonetheless, EF-Tu elicitor activity can easily be detected by *A. thaliana* cells in heat-killed bacterial soluble extracts[22]. We therefore compared the ability of bacterial extracts to elicit the production of ROS in *A. thaliana* leaves. Extracts from *Pectobacterium carotovorum* 193 (formerly known as *Erwinia carotovora* 193), *Pectobacterium atrosepticum* SCRI1043 (formerly known as *E. carotovora* ssp. *atroseptica* SCRI1043) and *Dickeya dadantii* 3534 (formerly known as *Erwinia chrysanthemi* 3534) all elicited ROS production in wild-type or *fls2* mutant *A. thaliana* leaves (**Fig. 1d**). In contrast, this response was completely abolished in *efr* and *fls2 efr* mutant leaves, revealing that the major PAMP in these extracts recognized by *A. thaliana* is EF-Tu. In conclusion, we were able to demonstrate EF-Tu eliciting activities in all phytopathogenic bacteria tested.

Next, we tested whether stable transformation of two solanaceous species with *EFR* confers responsiveness to EF-Tu. We have previously reported that *Agrobacterium*-mediated transient expression of EFR in *N. benthamiana* is sufficient to confer binding and responses to elf18 (ref. 4). We generated transgenic *N. benthamiana* plants expressing *EFR* under the control of its native promoter and selected homozygous plants carrying the *EFR* transgene for detailed phenotypic analysis (**Fig. 2a**). Whereas wild-type *N. benthamiana* plants were insensitive to elf18, transgenic EFR plants produced ROS in response to elf18 (**Fig. 2b**). In addition to triggering an oxidative burst, elf18 also induces the expression of defense-marker genes and seedling growth inhibition in *A. thaliana*[22]. Indeed, the expression of the PAMP-inducible genes *CYP71D20*, *FLS2*, *ACRE132* and *WRKY22*

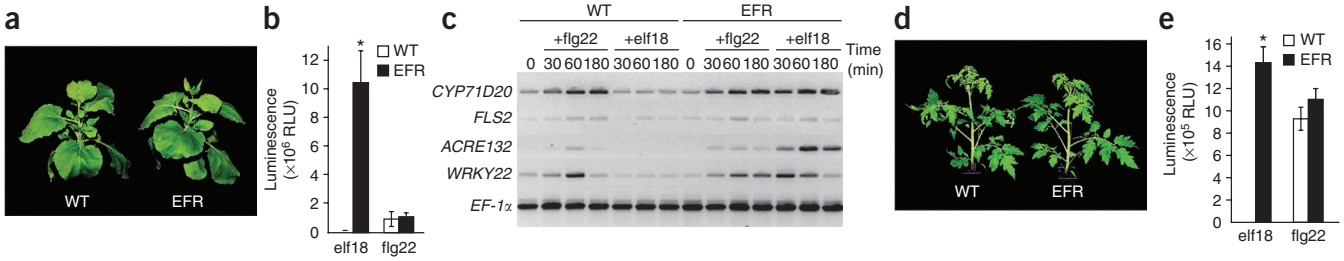


Figure 2 Transgenic expression of EFR in *N. benthamiana* and tomato confers elf18 responsiveness. (**a**) Four-week-old wild-type (left) and transgenic EFR (right) *N. benthamiana* plants. (**b**) Oxidative burst triggered by 100 nM elf18^{Ecoli} or flg22 in wild-type (white) or transgenic EFR (black) *N. benthamiana* leaf discs measured as RLU. Results are averages ± s.e.m. ($n = 12$). (**c**) Gene expression of marker genes determined by reverse transcriptase PCR. We treated wild-type and transgenic EFR *N. benthamiana* seedlings grown in axenic conditions with 100 nM elf18^{Ecoli} or flg22 for the times indicated. *EF-1α* is a housekeeping gene used a loading control. (**d**) Four-week-old wild-type (variety Moneymaker; left) and *EFR*-expressing transgenic tomato (right) plants. (**e**) Oxidative burst triggered by 100 nM elf18^{Ecoli} or flg22 in wild-type (variety Moneymaker; white) or transgenic EFR (black) tomato leaf discs measured as RLU. Results are averages ± s.e.m. ($n = 12$). *, $P < 0.05$ using Student's *t*-test. We repeated all experiments at least three times with similar results.

was induced by elf18 in the transgenic plants expressing *EFR*, but not in wild-type plants (**Fig. 2c**). In addition, the growth of the transgenic seedlings was inhibited *in vitro* in the presence of elf18, whereas wild-type seedlings grew normally (**Supplementary Fig. 1**).

Similarly, homozygous transgenic tomato lines (*S. lycopersicum* variety Moneymaker) expressing *EFR* under the control of the constitutive promoter *35S* also gained elf18 responsiveness (**Fig. 2d**,**e**). Together, these results show that stable expression of *EFR* in *N. benthamiana* and tomato confers responsiveness to elf18.

The introduction of a receptor that is normally absent from a plant species could potentially affect the function of a preexisting receptor by, for example, competing for common signaling partners. To test whether the transfer of EFR to solanaceous plants could affect the function of the related endogenous PRR FLS2, we compared responsiveness to flg22 in wild-type and transgenic *N. benthamiana* plants expressing *EFR*. flg22 induced similar levels of ROS production and defense-marker gene expression in wild-type and transgenic plants expressing *EFR* (**Fig. 2b**,**c**). Therefore the presence of EFR has no detrimental effects on endogenous levels of FLS2.

Next, we tested whether EF-Tu responsiveness in the transgenic plants expressing *EFR* is associated with increased disease resistance to adapted virulent bacteria. Inoculation of *N. benthamiana* with *P. syringae* pv. *syringae* (*Pss*) B728a, the causal agent of bacterial brown spot of bean, caused severe disease symptoms and substantial bacterial growth[23] (**Fig. 3a**). However, transgenic *N. benthamiana* plants expressing *EFR* developed less severe disease symptoms and allowed only 1.25% of the bacterial growth observed in wild-type leaves 4 d post-inoculation (**Fig. 3a**).

To assess whether EFR confers resistance to other *P. syringae* pathovars, we next infected transgenic *N. benthamiana* plants expressing *EFR* with *P. syringae* pv. *tabaci* (*Pta*) 11528, the causal agent of tobacco wildfire (also known as angular leaf spot). Leaves of transgenic plants expressing *EFR* showed only minor disease symptoms compared with wild-type plants (**Fig. 3b**). This corresponded with very weak growth of *Pta* 11528 bacteria in leaves from transgenic plants expressing *EFR*, reaching only 0.02% of the bacterial growth observed in wild-type leaves 4 d post-inoculation (**Fig. 3b**).

We have previously reported that loss of EF-Tu recognition in *A. thaliana* leads to increased susceptibility to *A. tumefaciens*, the causal agent of crown gall disease. This demonstrates that EFR contributes to resistance to this phytopathogen in *A. thaliana*[4]. *N. benthamiana* is normally very amenable to *Agrobacterium*-mediated transient expression (**Supplementary Fig. 2**). In contrast,

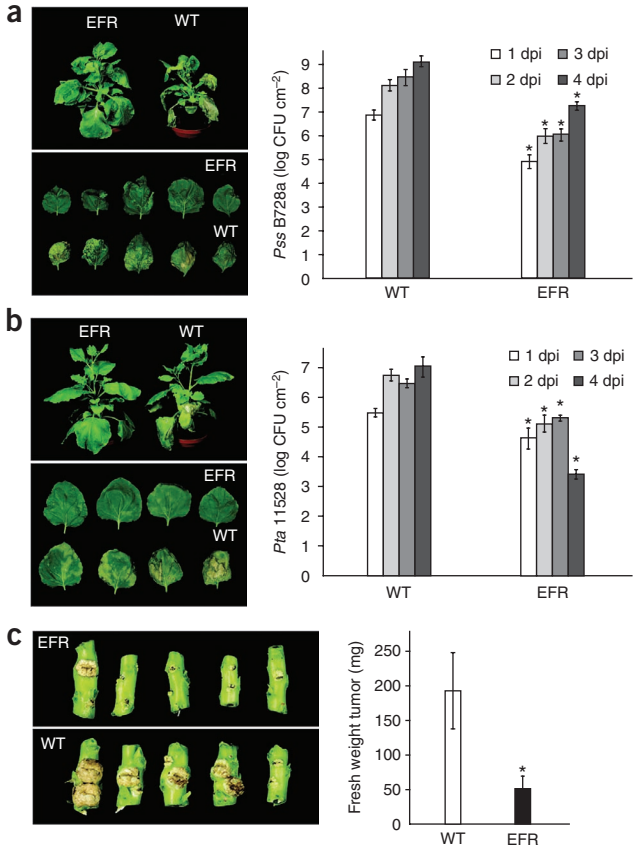


Figure 3 Transgenic expression of EFR in *N. benthamiana* confers broad-spectrum bacterial resistance. (**a**,**b**) Infection with *Pss* B728a (**a**) and *Pta* 11528 (**b**). We sprayed 4-week-old *N. benthamiana* plants with 10^8 colony-forming units (CFU) of bacteria per milliliter supplemented with 0.06% (vol/vol) Silwet-L77 and photographed them 6 days post-inoculation (dpi). Results are averages ± s.e.m. ($n = 4$). (**c**) Infection with *A. tumefaciens* A281. We stab-inoculated stems of 4-week-old *N. benthamiana* plants with bacteria that had been cultured on a plate for 2 d. We took pictures (left) and tumor fresh-weight measurements (right) at 3 weeks post-inoculation. Results are averages ± s.e.m. ($n = 16$). *, $P < 0.05$ using Student's *t* test. We repeated all experiments at least three times with similar results.

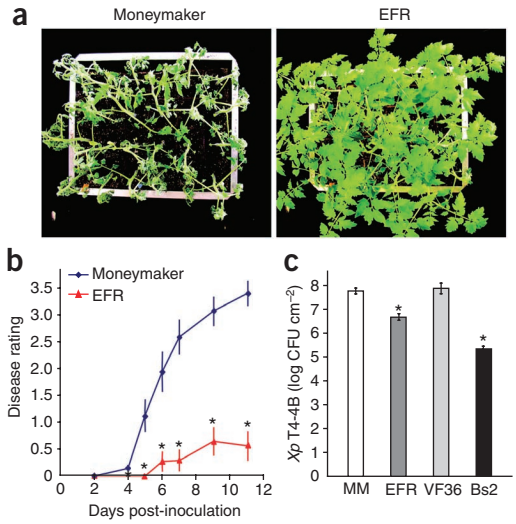


Figure 4 Transgenic expression of EFR in tomato confers broad-spectrum bacterial resistance. (**a**) Wild-type (variety Moneymaker; left) and transgenic EFR (right) tomato plants infected with *R. solanacearum* GMI1000. We drench-inoculated 4-week-old plants with 10^8 CFU ml^{-1} bacteria and photographed them 6 d after inoculation. (**b**) Disease scoring after infection with *R. solanacearum* GMI1000 in wild-type Moneymaker (blue) and transgenic EFR (red) tomato plants. Results are averages ± s.e.m. ($n = 24$). (**c**) Wild-type Moneymaker (MM) or VF36 and transgenic EFR- or Bs2-expressing tomato plants infected with *X. perforans* T4-4B. We dipped 6-week-old tomato plants in bacterial suspension (10^7 CFU ml^{-1} supplemented with 0.008% (vol/vol) Silwet-L77) and counted bacteria 14 d after inoculation. Results are averages ± s.e.m. ($n = 3$). *, $P < 0.05$ using Student's *t*-test. We repeated all experiments at least twice with similar results.

leaves from transgenic *N. benthamiana* plants expressing *EFR* were strongly impaired in the transient expression of a reporter transgene encoding the enzyme β-glucuronidase (GUS; **Supplementary Fig. 2**). Next, we tested whether EF-Tu perception restricts gall formation caused by a virulent tumorigenic *A. tumefaciens* strain. Notably, stab-inoculated stems of transgenic *N. benthamiana* plants expressing *EFR* developed tumors ~75% smaller than the ones present on wild-type stems (**Fig. 3c**).

Ralstonia solanacearum and *Xanthomonas perforans* (previously known as *X. axonopodis* pv. *vesicatoria*) are major pathogens of solanaceous plants, causing bacterial wilt and spot diseases, respectively. Whereas soil drenching with *R. solanacearum* led to massive wilting of wild-type tomato plants, transgenic plants expressing *EFR* showed drastically reduced wilting symptoms (**Fig. 4a**,**b**). Similarly, transgenic tomato plants expressing *EFR* were more resistant to *X. perforans* than were wild-type strains of tomato plants (**Fig. 4c** and **Supplementary Fig. 3**). Transgenic *EFR* expression conferred resistance to *X. perforans*, albeit to a lesser extent than transgenic expression of the *Bs2* gene from pepper (*Capsicum annuum*) (**Fig. 4c** and **Supplementary Fig. 3**). The *Bs2* gene encodes an R protein that confers strong field resistance via recognition of avrBs2, an effector widely conserved in *X. perforans* isolates[24]. Notably, the resistance conferred by EFR appears less efficient against *X. perforans* than against other bacteria tested (**Figs. 3** and **4**). Overall, these results demonstrate that EF-Tu perception by EFR in transgenic solanaceous plants enhances resistance to adapted foliar and vascular phytopathogenic bacteria.

One problem often associated with the heterologous overexpression of defense-related genes in plants is constitutive activation of defense responses that can lead to necrosis and/or reduced growth[6,7]. Yet sterile and soil-grown transgenic *N. benthamiana* and tomato plants expressing *EFR* did not show any constitutive activation of ROS production (data not shown) or defense-gene expression (**Fig. 2c** and **Supplementary Fig. 4**) and did not show any developmental or growth defects when grown in nonsterile soil over several generations (**Fig. 2a**,**d**). Consistent with the specificity of EFR for bacterial EF-Tu and the absence of constitutive activation of defense reactions, transgenic plants expressing *EFR* were not more resistant to the fungal pathogen *Verticillium dahliae* (**Supplementary Fig. 5**), which, like *R. solanacearum*, is a vascular phytopathogen. Therefore, transgenic expression of *EFR* does not seem to constitutively activate defense responses under laboratory conditions.

Genetic engineering of new PAMP recognition specificity across plant families offers several advantages over the current alternatives to improving resistance to phytopathogens. Using the plant's own immune system to combat plant diseases should reduce agrochemical inputs and their associated financial, health and environmental costs. In comparison to classical breeding for *R* gene–mediated resistance, transgenic approaches will allow rapid transfer of new PRRs into several elite varieties, as well as into crops that are not amenable to classical breeding (e.g., banana). In contrast to R proteins, a given PRR has the potential to confer resistance to a wide range of pathogens that carry the recognized PAMP. Here, for example, EFR conferred resistance to bacteria belonging to genera as diverse as *Pseudomonas*, *Agrobacterium*, *Xanthomonas* and *Ralstonia*. More importantly, gene-for-gene resistance is often rapidly broken down by evolving pathogens. Given the conserved and essential nature of PAMPs, it may be less likely for pathogens to evolve to evade recognition by PRRs; PAMP mutations are more likely to penalize fitness.

It should nonetheless be noted that rare examples of allelic variation in PAMP genes have been reported. These reflect a virulence strategy used by a small number of successful pathogens to avoid recognition[2,3]. In addition, some virulence effectors from the phytopathogenic bacterium *Pto* DC3000 are known suppressors of PTI in *A. thaliana* and tomato[25]. For example, AvrPto and AvrPtoB directly target PRRs, including FLS2 and EFR[17,26–28]. However, these effectors are not able to fully suppress PTI during infection, as mutations in PRRs enhance susceptibility to virulent bacteria[10,17]. Notably, our results (**Figs. 3** and **4**) suggest that effectors secreted during infection by *Pss* B728a, *Pta* 11528, *A. tumefaciens* A281, *R. solanacearum* GMI1000 or *X. perforans* T4-4B cannot fully suppress the immunity conferred by the transgenic expression of EFR in tomato and *N. benthamiana*. Moreover, because plants and their pathogens engage each other in a constant evolutionary arms race, the possibility of eventual recognition escape or PTI suppression by existing, newly acquired or evolved effectors cannot be discounted. However, the reduced pathogen populations resulting from PTI activation will decrease the probability of new virulent pathogen isolates arising and thus decrease the likelihood that resistance might break down[29].

Ultimately, our finding that PRRs can be transferred from one plant family to another provides a novel biotechnological approach to engineering disease resistance. We propose that combinations of several PRRs, as well as combinations of PRRs with R proteins recognizing widely distributed effectors (for example, Bs2) can be used to enable broad-spectrum disease resistance against multiple genera of plant pathogens with promising potential for durability under field conditions.

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 the John Innes Centre Horticultural Services for great plant care; D. Studholme (The Sainsbury Laboratory) for providing the EF-Tu sequence from *X. campestris* pv. *musacearum* 4381; S. Humphris, L. Pritchard, P. Birch and I. Toth (Scottish Crop Research Institute) for providing the EF-Tu sequences and extracts from *Dickeya dianthicola* 3534, *P. atrosepticum* 1043 and *P. carotovorum* 193; B. Vinatzer (Virginia Tech), J. Rathjen (The Sainsbury Laboratory) and S. Gelvin (Purdue University) for providing cultures of *Pss* B728a, *Pta* 11528 and *A. tumefaciens* A281, respectively; and S. Kamoun, E. Ward and J. Rathjen for their critical reading of the manuscript. This work was funded by the Two Blades Foundation and the Gatsby Charitable Foundation. B.P.H.J.T. is supported by a Vidi grant of the Research Council for Earth and Life Sciences of the Netherlands Organization for Scientific Research and by European Research Area Networks Plant Genomics. The Two Blades Foundation has filed a patent on behalf of inventors J.D.G.J. and C.Z. on the use of EFR to confer broad-spectrum disease resistance in plants.

AUTHOR CONTRIBUTIONS

S.L., A.R.-C., E.S., N.P., D.D., H.P.E. and G.R. performed experiments and analyzed data. M.S. generated the transgenic plants. B.S. and B.P.H.J.T. contributed ideas, conceived experiments and analyzed data. J.D.G.J. initiated the project and contributed ideas. C.Z. initiated the project, conceived, designed and performed experiments, analyzed data, obtained funding, and wrote the manuscript. All authors commented on the manuscript prior to submission.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

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

1. Jones, J.D. & Dangl, J.L. The plant immune system. *Nature* **444**, 323–329 (2006).
2. Zipfel, C. Early molecular events in PAMP-triggered immunity. *Curr. Opin. Plant Biol.* **12**, 414–420 (2009).
3. Boller, T. & Felix, G. A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annu. Rev. Plant Biol.* **60**, 379–406 (2009).
4. Zipfel, C. *et al.* Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts *Agrobacterium*-mediated transformation. *Cell* **125**, 749–760 (2006).
5. Reaping the benefits: science and the sustainable intensification of global agriculture (RS Policy document 11/09) (Royal Society, London, 2009).
6. Hammond-Kosack, K.E. & Parker, J.E. Deciphering plant-pathogen communication: fresh perspectives for molecular resistance breeding. *Curr. Opin. Biotechnol.* **14**, 177–193 (2003).
7. Gurr, S.J. & Rushton, P.J. Engineering plants with increased disease resistance: what are we going to express? *Trends Biotechnol.* **23**, 275–282 (2005).
8. Leach, J.E., Vera Cruz, C.M., Bai, J. & Leung, H. Pathogen fitness penalty as a predictor of durability of disease resistance genes. *Annu. Rev. Phytopathol.* **39**, 187–224 (2001).
9. McDonald, B.A. & Linde, C. Pathogen population genetics, evolutionary potential, and durable resistance. *Annu. Rev. Phytopathol.* **40**, 349–379 (2002).
10. Zipfel, C. *et al.* Bacterial disease resistance in *Arabidopsis* through flagellin perception. *Nature* **428**, 764–767 (2004).
11. Li, X. *et al.* Flagellin induces innate immunity in nonhost interactions that is suppressed by *Pseudomonas syringae* effectors. *Proc. Natl. Acad. Sci. USA* **102**, 12990–12995 (2005).
12. Hann, D.R. & Rathjen, J.P. Early events in the pathogenicity of *Pseudomonas syringae* on *Nicotiana benthamiana*. *Plant J.* **49**, 607–618 (2007).
13. de Torres, M. *et al.* *Pseudomonas syringae* effector AvrPtoB suppresses basal defence in *Arabidopsis*. *Plant J.* **47**, 368–382 (2006).
14. Wan, J. *et al.* A LysM receptor-like kinase plays a critical role in chitin signaling and fungal resistance in *Arabidopsis*. *Plant Cell* **20**, 471–481 (2008).
15. Miya, A. *et al.* CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **104**, 19613–19618 (2007).
16. Kaku, H. *et al.* Plant cells recognize chitin fragments for defense signaling through a plasma membrane receptor. *Proc. Natl. Acad. Sci. USA* **103**, 11086–11091 (2006).
17. Gimenez-Ibanez, S. *et al.* AvrPtoB Targets the LysM receptor kinase CERK1 to promote bacterial virulence on plants. *Curr. Biol.* **19**, 423–429 (2009).
18. Lee, S.-W. *et al.* A type I-secreted, sulfated peptide triggers XA21-mediated innate immunity. *Science* **326**, 850–853 (2009).
19. Gimenez-Ibanez, S., Ntoukakis, V. & Rathjen, J. The LysM receptor kinase CERK1 mediates bacterial perception in *Arabidopsis*. *Plant Signal. Behav.* **4**, 539–541 (2009).
20. Nekrasov, V. *et al.* Control of the pattern-recognition receptor EFR by an ER protein complex in plant immunity. *EMBO J.* **28**, 3428–3438 (2009).
21. Wang, G.L., Song, W.Y., Ruan, D.L., Sideris, S. & Ronald, P.C. The cloned gene, Xa21, confers resistance to multiple *Xanthomonas oryzae* pv *oryzae* isolates in transgenic plants. *Mol. Plant Microbe Interact.* **9**, 850–855 (1996).
22. Kunze, G. *et al.* The N terminus of bacterial elongation factor Tu elicits innate immunity in *Arabidopsis* plants. *Plant Cell* **16**, 3496–3507 (2004).
23. Vinatzer, B. *et al.* The type III effector repertoire of *Pseudomonas syringae* pv. *syringae* B728a and its role in survival and disease on host and non-host plants. *Mol. Microbiol.* **62**, 26–44 (2006).
24. Tai, T.H. *et al.* Expression of the Bs2 pepper gene confers resistance to bacterial spot disease in tomato. *Proc. Natl. Acad. Sci. USA* **96**, 14153–14158 (1999).
25. Cui, H., Xiang, T. & Zhou, J.M. Plant immunity: a lesson from pathogenic bacterial effector proteins. *Cell. Microbiol.* **11**, 1453–1461 (2009).
26. Xiang, T. *et al.* *Pseudomonas syringae* effector AvrPto blocks innate immunity by targeting receptor kinases. *Curr. Biol.* **18**, 74–80 (2008).
27. Shan, L. *et al.* Bacterial effectors target the common signalling partner BAK1 to disrupt multiple MAMP receptor-signaling complexes and impede plant immunity. *Cell Host Microbe* **4**, 17–27 (2008).
28. Gohre, V. *et al.* Plant pattern-recognition receptor FLS2 is directed for degradation by the bacterial ubiquitin ligase AvrPtoB. *Curr. Biol.* **18**, 1824–1832 (2008).
29. Brun, H. *et al.* Quantitative resistance increases the durability of qualitative resistance to *Leptosphaeria maculans* in *Brassica napus*. *New Phytol.* **185**, 285–299 (2010).

ONLINE METHODS

Plant growth. *A. thaliana* ecotype Columbia (Col-0) was the background for all mutant lines used in this study. We grew plants as one plant per pot at 20–21 °C with an 10-h photoperiod or on plates containing MS salts medium (Duchefa), 1% (wt/vol) sucrose and 1% (wt/vol) agar with a 16-h photoperiod. We grew *N. benthamiana* and tomato (*S. lycopersicum*) variety Moneymaker as one plant per pot at 22 °C with a 16-h photoperiod.

Generation of transgenic plants. We transferred the recombinant binary vectors *pGREENII-EFRp::EFR* and *pBIN19-35S::EFR* into *A. tumefaciens* strain Agl1 by electroporation. We generated transgenic *N. benthamiana* plants expressing *pGREENII-EFRp::EFR* as described[30], and transgenic tomato (*S. lycopersicum*) variety Moneymaker plants expressing *pBIN19-35S::EFR* essentially as described[31]. For *N. benthamiana*, we recovered 15 primary transformants after selection on phosphinothricin-containing plates. After transfer to soil, we collected leaf discs collected from individual primary transformant plants and tested them for gain of elf18 responsiveness in the luminal-based oxidative burst assay. Of 15 primary transformants, 9 gained elf18 responsiveness with similar levels. We randomly selected two of these lines and further selected homozygous progeny on the basis of segregation analyses on selection plates. The behavior of these lines was identical in terms of elf18 responsiveness as measured by oxidative burst and seedling growth inhibition. Therefore, we used only one line (no. 6-12-18) at the T4 stage for the final analysis reported in the manuscript.

For tomato, we recovered 20 primary transformants after selection on kanamycin-containing plant plates. After transfer to soil, we collected leaf discs from individual primary transformant plants and tested them for gain of elf18 responsiveness in the luminal-based oxidative burst assay. Of 20 primary transformants, 13 gained elf18 responsiveness with similar levels. In the next generations, we identified homozygous lines using segregation analysis on selection plates and quantitative real-time PCR[32]. We used two homozygous lines (no. P[1]-II7 and no. L[16]-I4) for physiological and pathological characterizations; these behaved similarly in all assays. The results presented in the manuscript are those of no. P[1]-II7.

Bioassays. We performed seedling growth inhibition and oxidative burst assays as previously described[25]. We ordered peptides from Peptron (http://www.peptron.com/).

To prepare bacterial extracts, we grew bacteria from a fresh culture on plates overnight in 50 ml of L medium (10 g l^{-1} bacto-peptone, 5 g l^{-1} yeast extract, 5 g l^{-1} sodium chloride, 1 g l^{-1} D-glucose, pH 7), pelleted them for 10 min, resuspended them in 2 ml sterile water, boiled them for 10 min at 95 °C, pelleted them for 10 min and collected the supernatant. We used 10 µl of bacterial extracts for the oxidative burst assays.

Sequence analysis. We retrieved EF-Tu sequences from UniProtKB (http://www.uniprot.org/help/uniprotkb) and generated the alignment using ClustalW2 (http://www.ebi.ac.uk/) and Boxshade (http://www.ch.embnet.org/) with default settings. We generated the consensus elf18 peptide using WebLogo (http://weblogo.berkeley.edu/).

Gene expression. We treated 2-week-old plants grown in liquid MS 1% (wt/vol) sucrose medium with 100 nM elf18 or flg22 for 0, 30, 60 or 180 min and froze them in liquid nitrogen. We extracted total RNA from seedlings with a RNeasy Plant Mini kit (Qiagen). We treated RNA samples with DNase Turbo DNA-free (Ambion), quantified with a Nanodrop spectrophotometer (Thermo Scientific) and reverse-transcribed 2 µg of total RNA into complementary DNA (cDNA) with SuperScript II reverse transcriptase (Invitrogen). We used 1 µl of cDNA in PCR under the following conditions: 95 °C, 2 min; (95 °C, 45 s; 58 °C, 45 s; 55 °C, 30 s; 72 °C, 1.5 min) × 25 cycles; 72 °C, 5 min.

We used the following primers: *CYP71D20*, 5′-AAGGTCCACCGCACCATGTCCTTAGAG-3′ and 5′-AAGAATTCCTTGCCCCTTGAGTACTTGC-3′; *FLS2*, 5′-AAGGATCCTGTGACTTGAAGCCTTCAA-3′ and 5′-AAGAATTCATTGGTAATTCATCAGCTCCTGTAA-3′; *ACRE132*, 5′-AAGGTCCAGCGAAGTCTCTGAGGGTGA-3′ and 5′- AAGAATTCCAATCCTAGCTCTGGCTCCTG-3′; *WRKY22*, 5′- AAGGTCCGGGATCTACATGCGGTGGT-3′ and 5′- AAGAATTCCGGGTCGGATCTATTTCG-3′; *PR1acidic*, 5′-TAGTCATGGGATTTGTTCTC-3′ and 5′-TCAGATCATACATCAAGCTG-3′; *EF-1α*, 5′-AAGGTCCAGTATGCCTGGGTGCTTGAC-3′ and 5′-AAGAATTCACAGGGACAGTTCCAATACCAC-3′.

We used the constitutively expressed *EF-1α* housekeeping gene as a loading control.

Disease assays. For infections with *Pss* B728a and *Pta* 11528, we sprayed four 4-week-old *N. benthamiana* plants with 10^8 CFU ml^{-1} bacteria in water supplemented with 0.06% (vol/vol) Silwet-L77. The bacterium inoculum was prepared from an overnight culture at 28 °C in King's B medium (20 g l^{-1} bacto-peptone, 1.5 g l^{-1} dipotassium hydrogen orthophosphate, 1% (vol/vol) glycerol, pH 7) started from a fresh culture on an agar plate. After spraying, we left plants uncovered for the duration of the experiment. We described bacterial populations by growth curve analysis as described[13].

For infections with *A. tumefaciens* A281, we stab-inoculated stems of 16 4-week-old *N. benthamiana* plants with a plastic tip that had been dipped into a fresh culture on an L-medium plate. We assessed tumor formation 3 weeks after inoculation by measuring the fresh weight of outgrowing tumors that were excised from the stem with a scalpel. We performed *Agrobacterium*-mediated transient expression with *pBIN19-35S::GUS:HA* as described[25].

For infections with *R. solanacearum* GMI1000, we transferred 4-week-old tomato plants into the inoculation facility (28 °C, 16-h photoperiod), grew bacteria in rich B medium (10 g l^{-1} bacto-peptone, 1 g l^{-1} yeast extract, 1 g l^{-1} casamino acids) and soil-drenched each tomato plant with 50 ml of 10^8 CFU ml^{-1}. We performed disease scoring daily using a visual index in which the numbers 1, 2, 3 and 4 correspond to 25%, 50%, 75% and 100% wilted leaves, respectively.

For infections with *X. perforans* T4-4B, we dipped 6-week-old tomato plants for 30 s into 10^7 CFU ml^{-1} bacteria in water supplemented with 0.008% (vol/vol) Silwett-L77. The bacterium inoculum was prepared from an overnight culture at 28 °C in NYGB medium (5 g l^{-1} bacto-peptone, 3 g l^{-1} yeast extract, 2% (vol/vol) glycerol, 1 mM magnesium chloride) started from a fresh culture on an agar plate. After dipping, we covered plants with plastic bags and returned them to the growth chamber for 3 d, then left them uncovered for the duration of the experiment. We described bacterial populations using growth curve analysis by grinding 1-cm^2 leaf samples in water and plating appropriate dilutions on NYGB supplemented with rifampicin 100 µg ml^{-1} to select for *X. perforans* T4-4B and cycloheximide 50 µg ml^{-1} to prevent fungal growth.

For infections with *V. dahliae*, we inoculated 10-day-old tomato plants or 4-week-old *N. benthamiana* plants with *V. dahliae* race JR2 (10^6 conidiospores per milliliter) by uprooting and subsequent root-dip inoculation, as described[33]. As a control, we mock-inoculated plants with water.

30. Fillatti, J.J., Kiser, J., Rose, R. & Comai, L. Efficient transfer of a glyphosate tolerance gene into tomato using a binary *Agrobacterium tumefaciens* vector. *Bio/Technology* **5**, 726–730 (1987).
31. Horsch, R.B. *et al.* A simple and general method of transferring genes into plants. *Science* **227**, 1229–1231 (1985).
32. German, A.M., Kandel-Kfir, M., Swarzberg, D., Matsevitz, T. & Granot, D. A rapid method for the analysis of zygosity in transgenic plants. *Plant Sci.* **164**, 183–187 (2003).
33. Fradin, E.F. *et al.* Genetic dissection of *Verticillium* wilt resistance mediated by tomato Ve1. *Plant Physiol.* **150**, 320–332 (2009).

doi:10.1038/nbt.1613

nature biotechnology

RESOURCE

High-resolution DNA analysis of human embryonic stem cell lines reveals culture-induced copy number changes and loss of heterozygosity

Elisa Närvä[1], Reija Autio[1,2], Nelly Rahkonen[1], Lingjia Kong[2], Neil Harrison[3], Danny Kitsberg[4], Lodovica Borghese[5], Joseph Itskovitz-Eldor[6], Omid Rasool[1], Petr Dvorak[7], Outi Hovatta[8], Timo Otonkoski[9,10], Timo Tuuri[9], Wei Cui[11], Oliver Brüstle[5], Duncan Baker[12], Edna Maltby[12], Harry D Moore[13], Nissim Benvenisty[14], Peter W Andrews[3], Olli Yli-Harja[2,15] & Riitta Lahesmaa[1]

Prolonged culture of human embryonic stem cells (hESCs) can lead to adaptation and the acquisition of chromosomal abnormalities, underscoring the need for rigorous genetic analysis of these cells. Here we report the highest-resolution study of hESCs to date using an Affymetrix SNP 6.0 array containing 906,600 probes for single nucleotide polymorphisms (SNPs) and 946,000 probes for copy number variations (CNVs). Analysis of 17 different hESC lines maintained in different laboratories identified 843 CNVs of 50 kb–3 Mb in size. We identified, on average, 24% of the loss of heterozygosity (LOH) sites and 66% of the CNVs changed in culture between early and late passages of the same lines. Thirty percent of the genes detected within CNV sites had altered expression compared to samples with normal copy number states, of which >44% were functionally linked to cancer. Furthermore, LOH of the q arm of chromosome 16, which has not been observed previously in hESCs, was detected.

Pluripotent hESCs are studied for potential applications in regenerative medicine because of their unique capacity to self-renew and to differentiate into any cell type. Although they can be grown indefinitely in culture, they commonly undergo adaptive changes during prolonged passaging *in vitro*. Such 'culture-adapted' cells tend to show increased growth rate, reduced apoptosis and karyotypic changes[1–5]. The genomic stability of hESCs is routinely monitored, and it is well established that they may acquire nonrandom gains of chromosomes, particularly chromosomes 12, 17 and X[5,6]. These changes show a striking similarity to those of germ cell tumors[3,5], suggesting that culture adaptation of hESCs may have parallels to tumor progression and emphasizing the need for thorough analysis of cells destined for clinical application.

The resolution of conventional karyotyping, or G-banding, is only 3–20 Mb. New DNA array–based methods, such as comparative genomic hybridization, increase the resolution from the Mb to the kb scale, enabling studies of CNVs[7] and LOH. CNVs are amplified or deleted regions ranging in size from intermediate (1–50 kb) to large (50 kb–3 Mb)[6,8,9] and are recognized as a major source of human genome variability. Specific recurrent CNVs are common in tumors[10,11]; particular tumor types have characteristic copy number patterns[12], and CNVs increase during tumor progression, influencing phenotypes and prognosis[11]. LOH is a well-known characteristic of many tumors resulting from the unmasking of recessive alleles and aberrant expression of imprinted genes[13]. It is possible that hESCs might exhibit uniparental disomy (a form of LOH) as observed in mouse ESCs (mESCs), such that both chromosomes are of maternal or paternal origin[3,14,15]. Detection of CNVs and LOH in hESCs could provide a sensitive measure of culture-induced changes.

The analytic methods used in previous studies of hESCs were not of sufficient resolution to detect all CNVs and LOH. The first comparative genomic hybridization study followed three lines over 30 passages using arrays with a resolution similar to that of conventional karyotyping and reported an abnormality of 46,X,idic(X)(q21)[16]. Another study compared early and late passages of nine lines using an Affymetrix array containing 115,000 probes[17]. The changes detected included an amplification of 17q, deletion of chromosome 13 and four large CNVs, one of which contained the *MYC* oncogene. A third study identified

[1]Turku Centre for Biotechnology, University of Turku and Åbo Akademi University, Turku, Finland. [2]Department of Signal Processing, Tampere University of Technology, Tampere, Finland. [3]Centre for Stem Cell Biology and the Department of Biomedical Science, University of Sheffield, Sheffield, UK. [4]Stem Cell Technologies Ltd., Jerusalem, Israel. [5]Institute of Reconstructive Neurobiology, Life & Brain Center, University of Bonn and Hertie Foundation, Bonn, Germany. [6]Faculty of Medicine, Technion-Israel Institute of Technology and Department of Obstetrics and Gynecology, Rambam Health Care Campus, Haifa, Israel. [7]Department of Biology, Faculty of Medicine, Masaryk University & Department of Molecular Embryology, Institute of Experimental Medicine, Academy of Sciences of the Czech Republic, Brno, Czech Republic. [8]Department CLINTEC, Karolinska Institutet, Karolinska University Hospital Huddinge, Stockholm, Sweden. [9]Program of Molecular Neurology, Biomedicum Stem Cell Center, University of Helsinki, Helsinki, Finland. [10]Children's Hospital, University of Helsinki, Helsinki, Finland. [11]Institute of Reproductive and Developmental Biology, Faculty of Medicine, Imperial College London, Hammersmith Campus, London, UK. [12]Sheffield Diagnostic Genetic Services, Sheffield Children's NHS Trust, Sheffield, UK. [13]Centre for Stem Cell Biology and the Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield, UK. [14]Stem Cell Unit, Department of Genetics, The Institute of Life Sciences, The Hebrew University, Jerusalem, Israel. [15]Institute for Systems Biology, Seattle, Washington, USA. Correspondence should be addressed to R.L. (riitta.lahesmaa@btk.fi) or E.N. (elisa.narva@btk.fi).

Received 24 June 2009; accepted 16 February 2010; published online 28 March 2010; doi:10.1038/nbt.1615

Table 1 HESC lines used in the study

hESC line	Passage (p)	Karyotype (G-banding)	Karyotyped at passage	Laboratory[a]
H7 (s14)	P30			P.W.A.
H7 (s14)	P38	46,XX[20]	P38	P.W.A.
H7 (s6)	P128			P.W.A.
H7 (s6)	P132	47,XX,+1,der(6)t(6;17)(q27;q1)[15] / 47,XX,+1,der(6)t(6;17)(q27;q1),i(20)(q10)[5]	P132	P.W.A.
H7 (s6)	P230			P.W.A.
H7 (s6)	P237	49,XXX,+add(1)(p3),der(6)t(6;17)(q27;q1),+20[30]	P237	P.W.A.
H7 (s6, teratoma)	P125			P.W.A.
H7 (s6, teratoma)	P127	47,XX,+add(1)(p1),der(6)t(6;17)(q27;q1),i(20)(q10)[30]	P127	P.W.A.
H7	P 91	46,XX[30]	P92	W.C.
H1	P 61	46,XY [12] / 46,XY,?dup(20)(q11.2q13.1)[21]	P63	W.C.
CCTL-10	P33	46,XY[30]	P24	P.D.
CCTL-12	P143	46,XX[30]	P143	P.D.
CCTL-14	P49	46,XX[30]	P40	P.D.
CCTL-14	P38	46,XX[30]	P40	P.D.
I6	P50	46,XY[30]	P41	N.B.
H9	P34	46,XX[30]	P33	N.B.
H9	P25	46,XX[20]	P27	R.L.
HS237	P135	46,X,idic(X)(q13)[30]	P135	R.L.
HS306	P35	46,XX[30]	P40	O.H.
I3 (I3.2)	P55	46,XX[30]	P50	N.B.
I3	P41	46,XX[30]	P41	O.B.
HS401	P53	46,XY[30]	P53	R.L.
HS293	P60			R.L.
HS293	P26	46,XY[30]	P37	O.H.
FES21	P51	46,XY,del(10)(q24)[1] / 46,XY[30]	P52	T.O.
FES22	P41	46, XY[11]	P42	T.O.
FES29	P37	46,XY, add(13)(p1)[1] / 46,XY[30]	P37	T.O.
FES61	P48	54, XY, +3,+5,+11,+12,+12,+16,+17,+20[15] / 54,XY,+3,+5,+11,+12,+12,+16,+16,+add(17)(q?23?),+20 [1] / 46, XY[15]	P50	T.O.
FES75	P19	47,XY,+12[2] / 46,XY[28]	P21	T.O.

First column describes the hESC line used. Further specification of the line is indicated inside brackets: (s14) = unadapted, (s6) = adapted, (teratoma) = samples were grown out of a teratoma in an immune-compromised mouse after the mouse had been injected with H7 (s6) cells[34]. (I3.2) = subclone of I3 created at P19.
[a]See author list for full names.

low-degree mosaicism of chromosome 13 trisomy for a short period during culture in one of the five lines, analyzed with normal metaphase comparative genomic hybridization target slides (Vysis) having a chromosome resolution of 400–550 bands[18]. More recently, 70 CNVs were detected in two hESC lines using Agilent arrays containing 236,000 probes[19], and, in another study, 22 abnormalities, ranging from 1.2 to 77.5 Mb, with a hotspot at 20q11.21, were identified in 17 lines with bacterial artificial chromosome/P1-plasmid artificial chromosome arrays[20].

Here we have analyzed 29 samples obtained at a range of passage numbers from 17 hESC lines of various origin. The analysis was performed with an Affymetrix SNP 6.0 array containing 906,600 probes for SNPs and 946,000 probes for CNVs. The array is suitable for detecting karyotype, CNV, LOH and SNP profiles. The intermarker distance of all the probes on the array is ≤0.7 kb, which considerably increases the genomic coverage and resolution compared with the previous platforms. The samples studied included karyotypically normal and abnormal samples as well as samples at low (<50) and high (>50) passage numbers. To study culture-induced changes, we included sample pairs of the same line grown in different laboratories as well as several samples of the H7 line during the adaptation process. We also examined whether the CNVs and large chromosomal changes that we identified affect gene expression by hybridizing RNA from nine samples to Human Exon 1.0 ST Arrays (Affymetrix).

RESULTS

Sample representation

Samples (**Table 1**) were provided by eight laboratories belonging to the ESTOOLS consortium (http://www.estools.eu/). Data were analyzed with the Affymetrix Genotyping Console 3.0.1, with a resolution configuration of 50 kb across the genome. Each hESC line had a unique SNP profile (**Supplementary Table 1**) as samples from individual lines maintained and cultured in different laboratories had identical SNP fingerprints, confirming that the line originated from the same individual.

A majority of CNVs contribute to amplifications

In all karyotypically normal chromosomes, we identified a total of 843 CNVs ranging in size from 50 kb to 3 Mb (**Supplementary Table 2**). In each of the samples, we identified on average 29 CNVs, with an average size of 221 kb and a median size of 133 kb. Based on the Toronto Database[8] (http://projects.tcag.ca/variation/), 79% of detected CNVs were known, 9% overlapped with known CNVs and 12% were novel. To compare these findings to the normal human genome, we analyzed 90 HapMap samples from Caucasians with identical analysis configurations (Online Methods) as with the hESC sample set (**Supplementary Table 3**). HapMap samples contained on average 26 CNVs per sample.

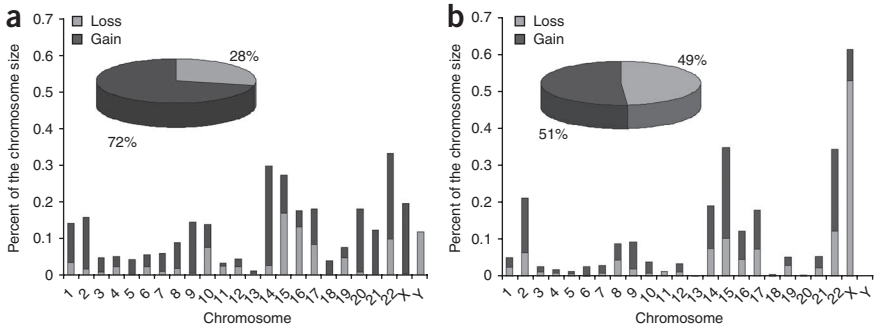


Figure 1 Amplifications contribute to majority of total genomic size affected by CNV in hESCs. (**a**,**b**) Average chromosomal distribution of 50 kb–3 Mb size CNVs in hESCs (**a**) and in Caucasian HapMap population (**b**). The majority (72%) of the total genomic size affected by CNVs found in hESCs corresponded to amplifications, whereas gains and losses were equally distributed in the HapMap samples. Chromosomal distribution differences between hESCs and HapMap were most prominent in chromosomes 10, 14, 20, X and Y.

Table 2 Large CNV changes (1–3 Mb in size) detected in hESC samples and genes within or overlapping these regions

Sample	Copy number state	Type	Chromosome	Start	End	Size (kb)	%CNV	Start	Name of variation	RefSeq genes on the area
HS306 P35	3	Gain	4	q22.1	q22.2	1,081	25	93332297	10054	*GRID2*
CCTL-12 P143	3	Gain	5	q14.2	q14.3	2,534	2	81717787	22770	*XRCC4*, *VCAN*, *HAPLN1*, *EDIL3*
I3.2 P55	3	Gain	10	q11.21	q11.22	1,203	100	46010225	0136	*PTPN20B*, *FRMPD2L2*, *FAM35B*, *SYT15*, *GPRIN2*, *PPYR1*, *ANXA8*, *ANXA8L1*
H7 s6 P128	1	Loss	10	q21.2	q21.3	1,288	15	63869872	30508	
H7 s6 P132	1	Loss	10	q21.2	q21.3	1,288	15	63869872	30508	*ZNF365*, *C10orf22* (also known as *ADO*), *EGR2*
H7 s6 Tera P125	1	Loss	10	q21.2	q21.3	1,288	15	63869872	30508	*NRBF2*, *JMJD1C*, *REEP3*
H7 s6 Tera P127	1	Loss	10	q21.2	q21.3	1,288	15	63869872	30508	
HS401 P53	1	Loss	15	q11.2	q11.2	1,009	100	18875309	0318	
H1 P61	1	Loss	15	q11.2	q11.2	1,243	100	18846092	0318	*HERC2P3*, *POTE15* (also known as *POTEB*)
H9 P25	3	Gain	15	q11.2	q11.2	1,357	100	18732853	0318	
H9 P34	3	Gain	15	q11.2	q11.2	1,434	100	18655531	0318	
HS237 P135	3	Gain	18	q21.32	q21.33	1,713	19	56145790	3171	*MC4R*, *CDH20*, *RNF152*
CCTL-14 P38	3	Gain	20	q11.21	q11.21	1,829	38	29298698	35916	*DEFB115/116/118/119/121/ 123/124*, *REM1*, *HM13*, *ID1*, *COX4I2*, ***BCL2L1***, *TPX2*, *MYLK2*, *FKHL18* (also known as *FOXS1*), *DUSP15*
CCTL-14 P49	3	Gain	20	q11.21	q11.21	1,831	38	29298698	35916	*TTLL9*, *PDRG1*, *XKR7*, *C20orf160*, *HCK*, *TM9SF4*, *PLAGL2*, *POFUT1*, *KIF3B*, *ASXL1*, *C20orf112*, *LOC149950*, *COMMD7*, ***DNMT3B***, *MAPRE1*, *SPAG4L* (also known as *SUN5*), *BPIL1*, *BPIL3*, *C20orf185*

These changes are below the detection limit of conventional karyotyping and were detected only with the array. %CNV, percent size of detected change overlapping location of the known genomic variation, if %CNV is 0 = novel CNV. Genes in boldface are involved with pluripotency and anti-apoptosis.

The average and median sizes were 232 kb and 127 kb, respectively, of which 80% were known, 10% overlapped with known CNVs and 10% were novel. Thus, the basic CNV statistics were similar in hESCs and the normal human genome. However, there were obvious differences in the pattern and distribution of the CNVs. These differences were most prominent in chromosomes 10, 14, 20, X and Y (**Fig. 1**). Strikingly, a clear majority (72%) of the total genomic size affected by CNVs in hESCs corresponded to amplifications, whereas in the HapMap samples gains and losses were equally distributed.

Fourteen of the CNVs detected were large, >1 Mb in size (**Table 2**). These were found only in the hESCs, with the exception of changes in 15q11.2, which were also detected in 30% of the HapMap samples. A change of particular interest was a 1,829-kb gain at 20q11.21 found in CCTL-14 passage (P)38/49. This region contains several genes, including *DNMT3B*, a known pluripotency-associated gene, and *BCL2L1*, which encodes the anti-apoptotic protein BCL-X. We validated the copy number gain in the gene area of *DNMT3B* by RT-PCR and also measured increased RNA production of DNMT3B in affected samples (**Supplementary Fig. 1a**).

LOH detected in 16q

All of the samples had heterozygous chromosomes except for the 16q arm of the hESC line FES21 (**Supplementary Fig. 2**). The karyotype of this line indicated a normal pair of chromosomes 16. However, these chromosomes had identical q arms based on the LOH profile.

CNV and LOH sites change in culture

To study whether CNV and LOH regions are vulnerable during prolonged culture, we compared samples of the same line at different passages (H9 P25/P34, CCTL-14 P38/P49, HS293 P26/P60, I3 P41/P55, H7 P30/P91). We reasoned that this analysis would detect only changes that had occurred during culture, excluding normal individual variation. We detected differences in CNV and LOH regions in all sample pairs studied (**Fig. 2a**). On average, 24% of the LOH sites and 66% of the CNVs had undergone changes between early and late passages. These values were considerably higher than the calculated false-positive estimate for CNVs (12.5%) (**Supplementary Table 4**). The number of LOH sites correlated positively with the number of passages between sample collections in four sample pairs. These data showed that new LOH sites were created at an average rate of 1.3 per passage. The LOH changes, which were on average 1,000 kb in size, were identified in all chromosomes except in chromosomes 21 and Y (**Supplementary Table 5**).

Next, we investigated whether the total genomic area affected by changes increases in culture. We concentrated on analyzing the

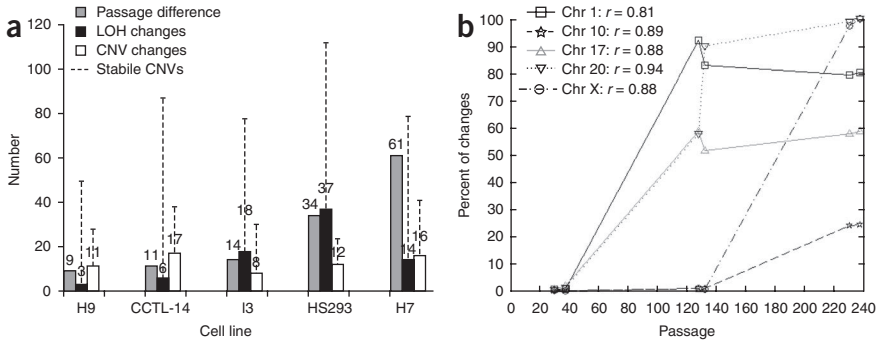


Figure 2 LOH and CNV regions change in culture. (**a**) The number of LOH, CNV and passages between sample collections in sample pairs (H9 P25/P34, CCTL-14 P38/P49, I3 P41/P55, HS293 P26/P60, H7 P30/P91). CNVs that remained stable during the culture are marked with dashed line. (**b**) The percentage of total genomic area changed plotted against the passages in culture shows clear correlation within chromosomes 1 (78%), 10 (89%), 17 (84%), 20 (90%) and X (88%) in H7 sample series, all $P < 0.05$. All seven samples are from the same hESC line H7 (P30, P38, P128, P132, P230 and P237). Large chromosomal changes in addition to CNVs were included in the analysis.

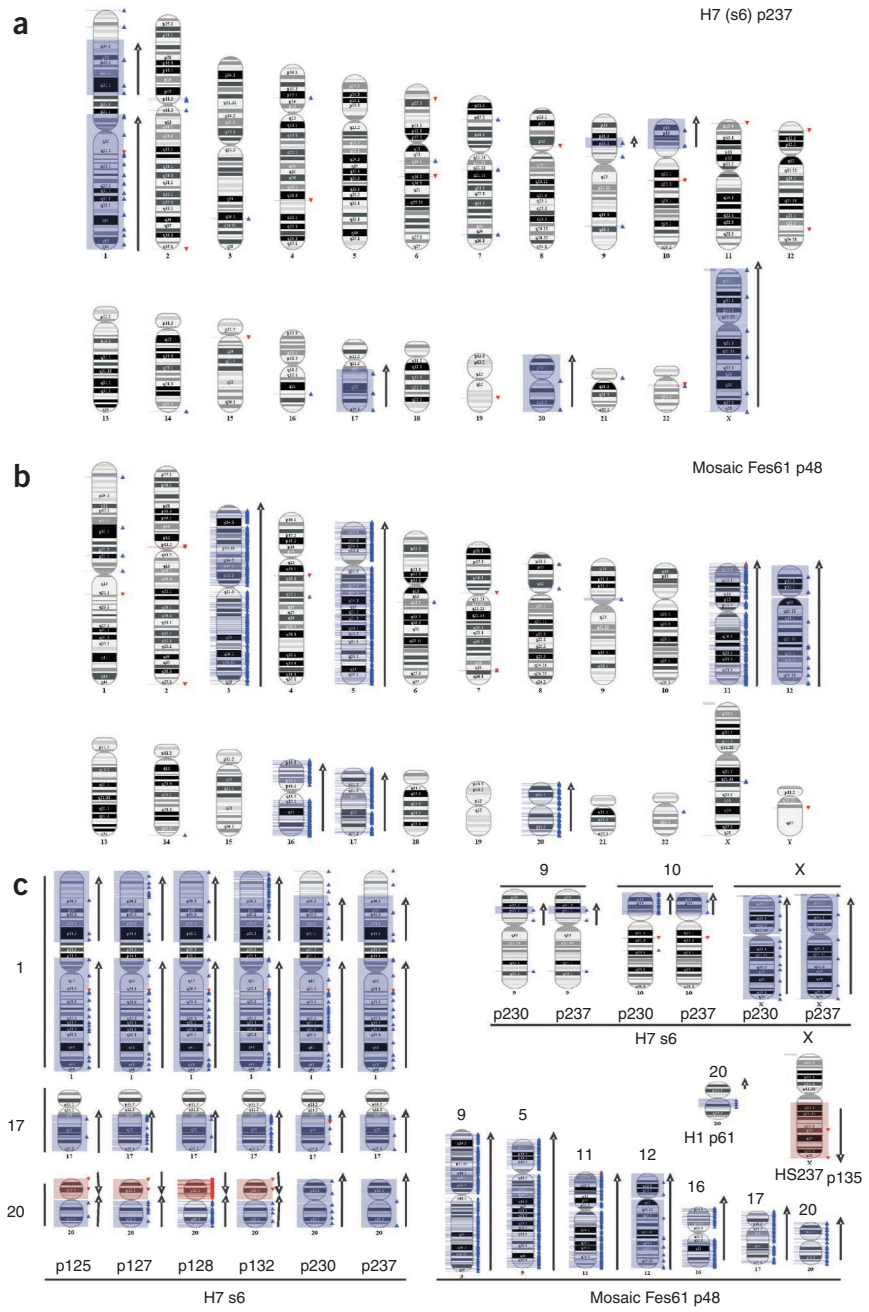


Figure 3 Chromosomal abnormalities detected. (**a**) The array karyotype of the sample H7 (s6) P237 shows deletions of extra abnormal chromosome 1 in 1p35 and in 1p terminus, as well as gains of 9p13–p21.2 and 10p11.2–p15, which were not seen by conventional karyotyping. (**b**) Mosaic karyotype of FES61, having an extra copy of chromosomes 3, 5, 11, 16, 17 and 20 and two extra copies of chromosome 12 in half of the cell population, was seen on the array karyoview as multiple CNVs in the chromosomes of the extra copy and total gain in the case of chromosome 12. (**c**) Summary of the large karyotype abnormalities detected. Gain, blue (↑); loss, red (↓). Each individual CNV is marked with a symbol: ▲, gain, ▼, loss.

unadapted (s14) P30/P38 and adapted (s6) middle P128/P132 and late P230/P237 samples of H7 cultured in similar conditions. The total size of large genomic areas and CNVs within each chromosome was studied in relation to the passage (**Supplementary Table 6**). We found a high correlation (from 0.83 to 0.97) in chromosomes 1, 10, 17, 20 and X, indicating that the actual chromosomal area with genomic changes increases in prolonged culture (**Fig. 2b**). Four randomly selected culture-induced CNVs were validated with RT-PCR. The CNVs (loss 10q21.2, 1,288 kb) and (gain 2q11.2, 213 kb) were present in all adapted samples but absent from unadapted samples, whereas (loss 6p23, 290 kb) and (gain 9q32, 896 kb) were present only in adapted late samples (**Supplementary Fig 3**).

Correlation between G banding

Abnormalities found with conventional karyotyping corresponded with the array data. For example, the array results of the line HS237 karyotyped as 46,X,idic(X)(q13) exhibited an 82,496 kb loss Xq13.1–q28. In addition, the arrays further clarified karyotype results. For instance, karyotype analysis showed that all H7 (s6) samples contained a structurally abnormal additional chromosome 1. The array indicated a gain of chromosome 1 except for p22.2–p21.1. Therefore, based on both methods the karyotype for chromosome 1 is +del(1)(p22.2p21.1). Notably, in H7 (s6) P230/P237 samples, besides detecting +del(1)(p22.2p21.1), the array also revealed a large deletion of 1p35 terminus in addition to gains of 9p13–p21.2 (12,038 kb) and 10p11.2–p15 (32,732 kb), which had not been detected by conventional karyotyping (**Fig. 3a**).

When conventional cytogenetics detected a mosaic karyotype, that is, 2 adapted cells among 30, the array could not detect any abnormalities. Conversely, if the sample contained a high level of mosaicism, the array detected multiple CNVs along affected chromosomes. For example, FES61 had a particularly complex karyotype, with one extra copy of chromosomes 3, 5, 11, 16, 17 and 20, and two of chromosome 12 in half of the population, with the other half being diploid. The array detected multiple CNVs in the chromosomes of one extra copy and a total gain in the case of chromosome 12 (**Fig. 3b**). Multiple CNVs created by a mosaic karyotype are just the sum result of two types of cell population on the array. The large chromosomal abnormalities detected are summarized in **Figure 3c**. These results suggest that the gain of 10p11.2–p15 in H7 (s6) was the result of a mosaic population at P230 that became further enriched by P237.

Shared variation between hESC lines

To study whether hESC lines share changes, we sorted out genes that showed CNVs in >25% of the samples. Seven amplified and two deleted regions were identified (**Table 3**). Many of the genes within these areas encoded immunoglobulin segments and olfactory receptors. However, many of these were also present in the 90 HapMap samples analyzed for comparison (**Supplementary Table 7** and **Table 3**). Notably, a deletion of a known tumor suppressor *HIC2* was found in eight samples. These deletions seemed to be culture induced because they were not

Table 3 Regions of variation shared by >25% of hESC samples

Average size (kb)	CNV%	Chromosome	Band	Biotype	Description	Genes	Gain in *n* samples	Loss in *n* samples
208	100	1	p36.13	Protein coding	Rootletin (ciliary rootlet coiled-coil protein)	*CROCC*	13	0
345	100	1	p36.33	Protein coding	Olfactory receptor	***OR4F5***	8	0
416	100	2	p11.2	*V* segment	Immunoglobulin κ light chain *V* gene segment	*IGKV1-5, IGKV4-1, IGKV2-24, IGKC*	27–29	0
124	100	7	q35	Protein coding	AP-4 complex subunit mu-1, seven transmembrane helix receptor	*AP4M1*, ***OR2A5***	7–13	2
442	100	14	q32.32	*C*/*V* segment	Immunoglobulin heavy chain *C*/*V* gene segments	***IGHM***, *IGHD, IGHV3-23, IGHG3, IGHV4-31*	15–29	0
578	100	15	q11.2	Protein coding	Olfactory receptor	*OR4N4, OR4M2*	2	9–14
267	100	21	p11.2	Protein coding	Putative tyrosine-protein phosphatase TPTE	*TPTE*	11	0
181	100	22	q11.22	*V* segment	Immunoglobulin λ light chain *V* gene segment	*IGLV2-23, IGLV2-18, IGLV2-11, IGLV2-14, IGLV3-25, IGLV3-22, IGLV3-21, IGLV3-16, IGLV3-12, IGLV3-19, IGLV4-3*	13–29	0
260	100	22	q11.21	Protein coding	Hypermethylated in cancer 2 protein (Hic-2) (Hic-3), **tumor suppressor**, putative phosphatidylinositol	***HIC2***, ***PI4KAP2***	0	8

Genes in boldface had <5% representation on (90) HapMap samples.

seen in the earlier passages of the lines affected. Some of the CNVs were specific for certain hESC lines. Four of these, 14q23.2, 305 kb, H9 P25/P34; 15q14, 103 kb, I3 P41/P55; 19q13.33, 345 kb, HS293 P15/P29; and 20q11.21, 1,829 kb, CCTL-14 P38/P49, were validated with RT-PCR (**Supplementary Fig. 4**).

Genes affected by CNVs

To investigate further which genes were affected by CNVs, we used the Ensembl (build 49) database[21] to find genes within CNV areas, resulting in a list of 354 genes (**Supplementary Table 8**). Notably, 77% of these corresponded to gene amplifications. We identified developmental genes *HOXA5,6,7,9,10,11* and *13*, which were affected by a 73-kb gain detected only in H7 (s6) P132 of the H7 samples, indicating that the change was culture associated. In addition, a gain of *DNMT3B* in both of the CCTL-14 samples was found, as mentioned above.

To identify genes associated with adaptation, we determined that 127 genes (**Supplementary Table 9**) had a different copy number in different passages of the same line (H9, CCTL-14, HS293, I3 and H7). Of these, 82% corresponded to amplifications and 19.1% were shared between different sample pairs. When these hits were compared to a list of oncogenes altered by CNVs (http://www.sanger.ac.uk/genetics/CGP/Census/)[10], within the area (155 kb gain 1q21.1) of the sample H7, we found a gene *PDE4DIP*, which is a known translocation gene in myeloproliferative disorder.

Genomic changes affect expression of genes

To study whether the CNVs and large chromosomal changes that we identified affect gene expression, we hybridized RNA from nine samples (FES21, 22, 29, 61, 75; H9 P25; H7 (s14) P38; H7 (s6) P132; H7 (s6) P237) to Human Exon 1.0 ST Arrays (Affymetrix). We integrated the copy number value with the gene expression values by computing a *P*-value for the association (**Supplementary Table 10**). With these settings, 29.9% of the genes had a significant (adjusted $p < 0.05$, fold change >2) increase in expression associated with an increase in copy number, whereas 41.6% of the copy number losses resulted in decreased expression. Next, we studied biological function related to these genes with Ingenuity Pathway analysis software (http://www.ingenuity.com/). The majority of the genes (44.4%) were linked to cancer; of these, 20.2% were associated with cell transformation and 14.3% with cell stage or division. Cancer types identified were gastrointestinal cancer, uterine tumor, ovarian cancer, non-Hodgkin lymphoma, myeloid leukemia, sarcoma, heart and pleura tumor, melanoma and central nervous system tumor.

To understand how culture-associated changes influence expression, we further studied the normal and adapted samples of H7. The majority of the changes were amplifications that increased expression. From the 1,121 gene amplifications detected only in adapted samples, 54.9% were identified already at P132 and the rest at P237. Thus, the number of changes influencing gene expression and the phenotype increased with prolonged time in culture. The most interesting gains found only in adapted H7 (s6) P237 sample were a cancer/testis-specific antigen *MAGEA4*, which was expressed over 17-fold, and *FGF13*, which was expressed over 2.5-fold at the RNA level compared to samples with normal copy number. In addition, the epigenetic regulator and cancer/testis gene *CTCFL* was expressed over tenfold in both adapted H7 (s6) samples P132 and P237. This gene has been shown to be co-expressed with *OCT-4* (also known as *POU5F1*) in hESCs at the protein level, transcribed in oocytes and downregulated in early cleavage stage embryos[22]. The gain of *MAGEA4* and *CTCFL* was validated with RT-PCR on the DNA and RNA levels (**Supplementary Fig. 1b**,**c**).

DISCUSSION

HESCs destined for therapeutic use should have a normal genetic composition. However, it is difficult to define 'normal' in this context as even the smallest change can have a substantial functional effect, for example, on the oncogenic potential of a cell. Our data show that genetic changes continue to increase during culture. Clearly, for clinical applications it will be important to minimize the time in culture. Our data did not allow us to define a safe cut-off passage. The average passage number of lines with a normal karyotype was 49.5 (median 41) compared with 110.8 (median 126) for lines with an abnormal karyotype. However, there were exceptions in both groups, as CCTL-12 had a normal karyotype at P142 and FES75 contained trisomy already at P19. In addition, the large 1–3 Mb changes affecting multiple genes were detected in both early and late passages. Some of the CNVs we identified were constitutional and were not acquired during culture. Ideally, the rest of the blastocyst used for hESC derivation or cells from very early passages should be stored as a standard procedure to facilitate identification of culture-induced changes.

We were able to confirm the identity of the different lines based on SNP profiles. This feature of the array can be used to verify the origin of different hESC lines. In addition, LOH was detected in the 16q arm. To our knowledge, LOH has not been reported previously in hESCs. The LOH of 16q is one of the most frequent somatic alterations in breast cancer[23] and occurs mainly in grade III tumors[24]. In addition, LOH of 16q has been identified in multiple myelomas and in prostate cancer[25,26]. We also showed that smaller LOH sites arise in culture. In mESCs, carcinogens are known to induce LOH[15], and a single insertion of the gene *neo* can undergo LOH as a result of selection pressure in culture, resulting in a duplicated *neo*-targeted locus[27]. Thus, it is not surprising that LOH can also occur in hESCs in culture.

We compared our data to earlier genomic studies of hESCs carried out with different array platforms. The HS237 line was reported to contain an aberrant X chromosome 46,X,idic(X)(q21) at p61 (ref. 16). In our analysis, HS237 had also deleted a part of the X chromosome at P135, that is, 46,X,idic(X)(q13), earlier karyotyped normal at P93. It is noteworthy that the same line grown in different laboratories undergoes a similar change, strengthening the conclusion that the change confers a selective advantage. Another study reported a deletion in chromosome 18 (ref. 17). We observed a 1,713 kb gain in this area that contains the genes *MC4R*, *CDH20* and *RNF152*. Recently, two studies reported recurrent genomic instability at 20q11.21 in multiple lines[20,28]. We also detected a 1,800 kb gain in this area in CCTL-14 samples.

Several mechanisms that may contribute to the genomic instability of hESCs have been identified. HESCs have an abnormal DNA repair system in that the mitotic spindle assembly checkpoint is functional but does not initiate apoptosis as it does in somatic cells[29]. In addition, hESCs downregulate the mismatch repair system when cultured in hypoxic conditions[30]. Furthermore, hESCs can accommodate LINE-1 retrotransposition, which could promote genomic fluidity[31].

Of the 354 genes affected by CNVs, 77% were located on ampliareas. Considering only CNVs that were culture induced, amplifications were observed in 82% of the affected genes. The greater proportion of amplifications in culture-induced CNVs might be explained by the process of adaptive amplification, in which amplification occurs as a part of the general stress response with which cells adjust to culture conditions[32]. CNVs can affect the phenotype of cells by altering coding and regulatory sequences or by amplifying or deleting gene copies. We found that CNVs changed the expression level of 30 % of the genes overlapping CNVs. Notably, >44% of genes whose expression was altered by CNVs were associated with cancer, emphasizing the importance of careful monitoring of hESCs to be used for clinical applications.

In the future, it will be of interest to study whether CNVs influence the varying differentiation potential of hESC lines[33]. In addition, high-resolution genomic analysis could be used to elucidate possible rearrangement in the reprogramming process of induced pluripotent cells. Furthermore, advances in sequencing technology are expected to overcome limitations in analytic resolution, enabling identification of minor genomic changes that will facilitate understanding of the adaptation, pluripotency, differentiation and tumorigenicity of hESCs.

METHODS

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

Accession codes. NCBI Gene Expression Omnibus: GSE15097.

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

ACKNOWLEDGMENTS

We are grateful to everyone who has taken care of sample collection and handling: T. Golan-Lev, A. Urrutikoetxea-Uriguen, S. Haupt, P. Koch, I. Laufenberg, B. Ley, A. Hampl, M. Vodinska, K. Koudelkova, S. Ström, F. Holm, A.-M. Strömberg, C. Olsson, M. Mikkola, S. Vuoristo, P. Junni and M. Hakkarainen. We especially acknowledge M. Linja, T. Heinonen and the Finnish DNA Microarray Centre for their excellent technical assistance. We acknowledge the Turku Graduate School of Biomedical Sciences. This study is supported by funding for the ESTOOLS consortium under the Sixth Research Framework Programme of the European Union, Juvenile Diabetes Research Foundation, The Academy of Finland and the Finnish Cancer Organizations, The Improving Outcomes Guidance Trust, The Ministry of Education, Youth, and Sport of the Czech Republic, Ida Montin Foundation, The Academy of Finland, projects no. 129657 (Finnish Centre of Excellence program 2006-11) and no. 134117 and the Medical Research Council, UK.

AUTHOR CONTRIBUTIONS

E.N., R.A., N.B., P.W.A., O.Y.-H. and R.L. designed the experiments, E.N. and R.L. were responsible for the coordination of the project and microarray experiments. R.A., E.N. and O.Y.-H. were responsible for data analysis, integration and statistical analysis. N.R. performed RNA extractions. L.K. built the gene annotation list of genes overlapping CNVs. D.B. performed conventional karyotyping. E.N. and N.R. performed copy-number state validations with RT-PCR. J.I.-E. provided I3 and I6 lines for the study. P.D., O.H., T.O., T.T., N.B., W.C., O.B., E.M., H.D.M., P.W.A., O.Y.-H. and R.L. provided the samples and coordinated the project in their groups. E.N., R.A., N.R., L.K., N.H., D.K., L.B., J.I.-E., O.R., P.D., O.H., T.O., T.T., N.B., W.C., O.B., D.B., E.M., H.D.M., P.W.A., O.Y.-H. and R.L. contributed to writing the paper.

COMPETING FINANCIAL INTERESTS

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. Draper, J.S., Moore, H.D., Ruban, L.N., Gokhale, P.J. & Andrews, P.W. Culture and characterization of human embryonic stem cells. *Stem Cells Dev.* **13**, 325–336 (2004).
2. Draper, J.S. *et al.* Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells. *Nat. Biotechnol.* **22**, 53–54 (2004).
3. Hanson, C. & Caisander, G. Human embryonic stem cells and chromosome stability. *APMIS* **113**, 751–755 (2005).
4. Enver, T. *et al.* Cellular differentiation hierarchies in normal and culture-adapted human embryonic stem cells. *Hum. Mol. Genet.* **14**, 3129–3140 (2005).
5. Baker, D.E. *et al.* Adaptation to culture of human embryonic stem cells and oncogenesis in vivo. *Nat. Biotechnol.* **25**, 207–215 (2007).
6. Redon, R. *et al.* Global variation in copy number in the human genome. *Nature* **444**, 444–454 (2006).
7. Feuk, L., Carson, A.R. & Scherer, S.W. Structural variation in the human genome. *Nat. Rev. Genet.* **7**, 85–97 (2006).
8. Iafrate, A.J. *et al.* Detection of large-scale variation in the human genome. *Nat. Genet.* **36**, 949–951 (2004).
9. Sebat, J. *et al.* Large-scale copy number polymorphism in the human genome. *Science* **305**, 525–528 (2004).
10. Futreal, P.A. *et al.* A census of human cancer genes. *Nat. Rev. Cancer* **4**, 177–183 (2004).
11. Kallioniemi, A. CGH microarrays and cancer. *Curr. Opin. Biotechnol.* **19**, 36–40 (2008).
12. Jong, K. *et al.* Cross-platform array comparative genomic hybridization meta-analysis separates hematopoietic and mesenchymal from epithelial tumors. *Oncogene* **26**, 1499–1506 (2007).
13. Zheng, H.T., Peng, Z.H., Li, S. & He, L. Loss of heterozygosity analyzed by single nucleotide polymorphism array in cancer. *World J. Gastroenterol.* **11**, 6740–6744 (2005).
14. Cervantes, R.B., Stringer, J.R., Shao, C., Tischfield, J.A. & Stambrook, P.J. Embryonic stem cells and somatic cells differ in mutation frequency and type. *Proc. Natl. Acad. Sci. USA* **99**, 3586–3590 (2002).
15. Donahue, S.L., Lin, Q., Cao, S. & Ruley, H.E. Carcinogens induce genome-wide loss of heterozygosity in normal stem cells without persistent chromosomal instability. *Proc. Natl. Acad. Sci. USA* **103**, 11642–11646 (2006).
16. Inzunza, J. *et al.* Comparative genomic hybridization and karyotyping of human embryonic stem cells reveals the occurrence of an isodicentric X chromosome after long-term cultivation. *Mol. Hum. Reprod.* **10**, 461–466 (2004).
17. Maitra, A. *et al.* Genomic alterations in cultured human embryonic stem cells. *Nat. Genet.* **37**, 1099–1103 (2005).
18. Caisander, G. *et al.* Chromosomal integrity maintained in five human embryonic stem cell lines after prolonged in vitro culture. *Chromosome Res.* **14**, 131–137 (2006).

19. Wu, H. *et al.* Copy number variant analysis of human embryonic stem cells. *Stem Cells* **26**, 1484–1489 (2008).
20. Spits, C. *et al.* Recurrent chromosomal abnormalities in human embryonic stem cells. *Nat. Biotechnol.* **12**, 1361–1363 (2008).
21. Hubbard, T.J. *et al.* Ensembl 2007. *Nucleic Acids Res.* **35**, D610–D617 (2007).
22. Monk, M., Hitchins, M. & Hawes, S. Differential expression of the embryo/cancer gene ECSA(DPPA2), the cancer/testis gene BORIS and the pluripotency structural gene OCT4, in human preimplantation development. *Mol. Hum. Reprod.* **14**, 347–355 (2008).
23. Lindblom, A., Rotstein, S., Skoog, L., Nordenskjold, M. & Larsson, C. Deletions on chromosome 16 in primary familial breast carcinomas are associated with development of distant metastases. *Cancer Res.* **53**, 3707–3711 (1993).
24. Cleton-Jansen, A.M. *et al.* Different mechanisms of chromosome 16 loss of heterozygosity in well- versus poorly differentiated ductal breast cancer. *Genes Chromosom. Cancer* **41**, 109–116 (2004).
25. Carter, B.S. *et al.* Allelic loss of chromosomes 16q and 10q in human prostate cancer. *Proc. Natl. Acad. Sci. USA* **87**, 8751–8755 (1990).
26. Jenner, M.W. *et al.* Gene mapping and expression analysis of 16q loss of heterozygosity identifies WWOX and CYLD as being important in determining clinical outcome in multiple myeloma. *Blood* **110**, 3291–3300 (2007).
27. Mortensen, R.M., Conner, D.A., Chao, S., Geisterfer-Lowrance, A.A. & Seidman, J.G. Production of homozygous mutant ES cells with a single targeting construct. *Mol. Cell. Biol.* **12**, 2391–2395 (1992).
28. Lefort, N. *et al.* Human embryonic stem cells reveal recurrent genomic instability at 20q11.21. *Nat. Biotechnol.* **26**, 1364–1366 (2008).
29. Mantel, C. *et al.* Checkpoint-apoptosis uncoupling in human and mouse embryonic stem cells: a source of karyotpic instability. *Blood* **109**, 4518–4527 (2007).
30. Rodriguez-Jimenez, F.J., Moreno-Manzano, V., Lucas-Dominguez, R. & Sanchez-Puelles, J.M. Hypoxia causes downregulation of mismatch repair system and genomic instability in stem cells. *Stem Cells* **26**, 2052–2062 (2008).
31. Garcia-Perez, J.L. *et al.* LINE-1 retrotransposition in human embryonic stem cells. *Hum. Mol. Genet.* **16**, 1569–1577 (2007).
32. Hastings, P.J. Adaptive amplification. *Crit. Rev. Biochem. Mol. Biol.* **42**, 271–283 (2007).
33. Osafune, K. *et al.* Marked differences in differentiation propensity among human embryonic stem cell lines. *Nat. Biotechnol.* **26**, 313–315 (2008).
34. Andrews, P.W. *et al.* Embryonic stem (ES) cells and embryonal carcinoma (EC) cells: opposite sides of the same coin. *Biochem. Soc. Trans.* **33**, 1526–1530 (2005).

ONLINE METHODS

Sample handling. Each hESC line isolated from the inner cell mass of *in vitro* fertilized genetically unique blastocyst was grown in each collaboration laboratory. Samples (**Table 1**) containing 1–2 million cells were harvested in the collaborating laboratory and sent frozen. Most of the samples were karyotyped also by conventional karyotyping. The culture technique and the media composition varied in different laboratories (**Supplementary Table 11**). Genomic DNA was extracted using QIAamp DNA Mini Kit (Qiagen). Concentration and quality of the samples was measured with spectrophotometer (Nanodrop, Thermo Scientific) and gel electrophoresis using Reference DNA as a control. All 29 samples were hybridized in the Finnish DNA Microarray Centre, at the Turku Centre for Biotechnology, using Genome-Wide Human SNP Nsp/Sty 6.0 protocol and SNP 6.0 arrays (Affymetrix).

For expression analysis, RNA was isolated using RNeasy Kit (Qiagen). To eliminate DNA from RNA samples DNase I (Qiagen) digestion was performed. Concentration of the samples was measured with Nanodrop. The selected nine samples (FES21, FES22, FES29, FES61, FES75, H9 (P25), H7 (s14) P38, H7 (s6) P132, H7 (s6) P237) were hybridized in the Finnish DNA Microarray Centre, at the Turku Centre for Biotechnology accordingly to manufacturer's protocol and hybridized on GeneChip Human Exon 1.0 ST Arrays (Affymetrix).

SNP 6.0 analysis. Data were analyzed using Affymetrix Genotyping Console 3.0.1 and Birdseed v2-algorithm. Samples were normalized against 40 International HapMap samples[35], which were also hybridized in-house to decrease technical variation. Sample codes for HapMap samples used are presented in the **Supplementary Table 1**. For the copy number analysis, we used regional GC correction and required ten markers to be found within the changed region and the size of the region to be at least 50 kb. All the arrays passed quality control requirements having contrast QC (Quality control) and MAPD (Median absolute pairwise difference) values within boundaries (**Supplementary Table 12**). Genotyping Console Browser (Affymetrix) was used to illustrate changes detected.

CNVs, in which the average distribution between markers was >20 kb, were considered as false positive in addition to CNVs affecting Y chromosome in female samples and excluded from the analysis. The false-positive estimate was studied by hybridizing three different HapMap samples in four replicates (**Supplementary Table 4**). By using identical analysis settings as for the main data, we found that on average 62% of CNVs were detected in all four replicates, 10.9% in three, 14.6% in two and 12.5% only in one of the replicates. These values are analogous with an earlier study[6]. We also analyzed all the CNV values across the genome of the sets of replicates, and on average 99.95% of the regions of all the replicates returned the same CNV value, either gained, normal or lost.

Ensembl (build 49) database was used to find the genes within the CNV areas[21]. The genes were further linked to HUGO Gene Nomenclature Committee gene symbols[36]. To compare an hESC CNV profile to a normal human genome, we analyzed 90 additional CEPH samples (Caucasians, Utah residents with Northern and Western European ancestry from the Centre d'Etude du Polymorphisme Humain collection) from the International HapMap Project (http://www.hapmap.org/) with identical settings to our own. The CEPH samples were chosen because they represent best the same sample origin as the hESC lines used in the study.

Exon array analysis. The probe values of the array were directly linked to Ensembl genes (build 49)[21] using alternative CDF-files, version 11 (ref. 37). We used the aroma.affymetrix package[38] in analyzing the gene values of the expression measurements, and used RMA[39] for pre-processing the Exon array values.

Integration of genomic changes and gene expression. To find the genes of which CNV is associated with increased or decreased gene expression level, we performed an integration analysis. First, we labeled the gene values into two groups; 'gain' and 'no gain'. For each gene, we computed a weight value

$$w_G = \frac{(m_{G1} - m_{G0})}{(\sigma_{G1} + \sigma_{G0})},$$

where G is the gene in question, m_{G1} and σ_{G1} denote the mean value and s.d. of the gene expression values of the samples, in which the gene was found to be gained, and m_{G0} and σ_{G0} the mean and s.d. of the samples, in which the gain was not detected[40]. To associate the lost copy number values with the low gene expression values, we labeled the genes into groups 'loss' and 'no loss', respectively, and computed the weight value for the association between a loss in copy number and a low gene expression value.

Second, we obtained a *P*-value for the weight value of each gene by performing 10,000 permutations[40]. Thus, we could identify genes with significant association between copy number and gene expression value. Third, the resulting *P*-values were adjusted with Benjamini Hochberg's multiple comparison method[41]. All the associations with over a twofold change between the mean values of the expression levels of groups 'gain' and 'no gain', or 'loss' and 'no loss' and the adjusted *P*-value >0.05 were considered to be significant[42].

Real time quantitative RT-PCR validation of the copy number states. To validate genomic copy number states, we used DNA from the original samples as a template. For the RNA analysis the RNA was isolated using RNeasy Kit. To eliminate genomic DNA from RNA samples, we included DNase I digestion in the column. Concentration of the samples was measured with Nanodrop. A second round of DNAse treatment was carried out for 500 ng of total RNA with DNase I Amplification Grade (Invitrogen). To verify that no genomic DNA was present, we performed negative RT-PCR control by measuring levels of the housekeeping gene EF1α. Subsequently, cDNA was prepared using a Superscript II kit (GIBCO). Gene expression levels were measured using the 7900HT Fast Real-Time PCR System (Applied Biosystems) using 2 μl of the template in 10 μl reaction volume. The primers and probes used were designed using Universal ProbeLibrary Assay Design Center (Roche). The primers designed for the analysis were first validated to respond by standard curve validation. All measurements were performed in duplicate in two separate runs and repeated if necessary to produce four Ct (threshold cycle) values for each gene where s.d. < 0.5. ΔCt for each gene was calculated ΔCt = Ct(gene) − Ct(GAPDH). The average results of the samples shearing gain (CN 3) or loss (CN 1) was compared to the samples of normal CN state (CN 2) for each gene studied. CN was counted real if the difference measured was in range of expected difference, 0.5 ΔCt for CN state 3 and 1 ΔCt for CN state 1. The two-tailed *t*-test was counted for each result and required to be under 0.05 (*), 0.01 (**) or 0.001 (***). Copy number states including loss and gains and size varying from 103 kb to chromosomal changes were selected for validation. 92% of the CNV selected for validation were verified with RT-PCR analysis.

Primers 5′–3′:

GAPDH: ACACCCACTCCTCCACCTTT, TGACAAAGTGGTCGTTGAGG, probe:45

DNMT3B: TGTAATCCAGTGATGATTGATGC, GGTAGGTTGCCCCAGAAGTAT, probe:84

RHOJ: GATGAGCTACGCCAACGAC, GCATAGTGGTCAAACACAGTGG, probe:6

CTCFL: GTGAGAAGCCTCACCTGTGTC, CGCAGCAGAGTGACCGTA, probe:13

EGR2: GGGTGTGTGCACCATGTC, GGTGGCGGAGAGTACAGGT, probe:85

MAGEA4: CCAATGAGGGTTCCAGCA, AACAAGGACTCTGCGTCAGG, probe:35

ZNF613: GGCAACCTCCTTATTCATCG, AGCCTTTCCCACATTCATTG, probe:47

ID1: CCAGAACCGCAAGGTGAG, GGTCCCTGATGTAGTCGATGA, probe:39

REV1: CCGGGAACAAGTAGAGCAAG, TTTTTGTCGCCATGTGACTC, probe:56

JARID2: TTCGCTCAGGAAAAAGAAGTG, AGTCATTGAGGACGCCTTTG, probe:63

TNFSF15: ACAGCCAGTGTGGAAATGCT, CCAGGCAGCAGGTGAGAG, probe:68

JMJD1C: GCAAACTGGGGAATCCTTTT, TTCTCGACACTTTTGTAAATTAGGC, probe:18

GOLGA8B: TGGCTTATTTCCGAGGAATG, CAAATGCTCTAAGCTAGGAAAGGT, probe: 76

RNA

EF1α: CTGAACCATCCAGGCCAAAT, GCCGTGTGGCAATCCAAT, probe: 6 (FAM)-AGCGCCGGCTATGCCCCTG-(TAMRA)

35. The International HapMap Consortium The international HapMap project. *Nature* **426**, 789–796 (2003).
36. Eyre, T.A. *et al.* The HUGO gene nomenclature database, 2006 updates. *Nucleic Acids Res.* **34**, D319–D321 (2006).

doi:10.1038/nbt.1615

37. Dai, M. *et al.* Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res.* **33**, e175 (2005).
38. Bengtsson, H., Simpson, K., Bullard, J. & Hansen, K.. *Aroma.Affymetrix: A Generic Framework In R For Analyzing Small To Very Large Affymetrix Data Sets In Bounded Memory. Technical report 745.* (Department of Statistics, University of California, Berkeley, 2008).
39. Bolstad, B.M., Irizarry, R.A., Astrand, M. & Speed, T.P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185–193 (2003).
40. Hautaniemi, S. *et al.* A strategy for identifying putative causes of gene expression variation in human cancers. *J. Franklin Inst.* **341**, 77–88 (2004).
41. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc., B* **57**, 289–300 (1995).
42. Jarvinen, A.K. *et al.* Identification of target genes in laryngeal squamous cell carcinoma by high-resolution copy number and gene expression microarray analyses. *Oncogene* **25**, 6997–7008 (2006).

Building a sustainable career in science

Aakanksha Singhvi & Pallavi Sachdev

Establishing a successful academic career in the age of 'big science'.

The word "sustainability" derives from the Latin word *sustinere* (*tenere*, to hold; *sus-*, up), to support or endure. A principal concern of postdocs is our path to initiating and sustaining a fruitful career. Whereas approximately 79% of postdocs start out aiming for a tenure-track academic position, only about 30% end up with one[1]. What does it take for postdocs to decide whether or not they want an academic job? What is a successful strategy for getting and sustaining one?

This was the subject of a roundtable discussion with David Roth, chairman of the department of pathology at New York University, at the fourth annual Rockefeller University postdoctoral retreat. Afterward, we polled Rockefeller postdocs and faculty on issues that were raised at the discussion. Here, we compile the opinions voiced at the retreat and in the surveys.

Defining "sustainability" in science

Our discussion title starting out was vague and open-ended on purpose. We wanted postdocs and faculty to tell us what concerned them most when thinking of their path to getting and sustaining a scientific career. Our own thoughts included getting a tenure-track position; sustaining funding; and becoming good mentors. In addition, to some, this meant balancing cost-effectiveness with creativity; sustaining a flow of creative ideas; ensuring that one's data upholds validity over time, and doing science without compromises. One Rockefeller faculty member defined it as the ability to conduct efficient research with effective reagents, tools and manpower in a cost-effective manner. Roth put foremost the ability to attract good students to one's lab. But although a good career in science needs all of the above, many of us worry as much about the ability to sustain scientific temper and creativity as about funding and job-seeking. In the next few sections, we summarize these various facets of building a scientific career.

Aakanksha Singhvi and Pallavi Sachdev are at Rockefeller University, New York, New York, USA.
e-mail: asinghvi@rockefeller.edu, sachdep@rockefeller.edu

Sustainability in developing a scientific career. The median age for landing a faculty position is 38, and for one's first R01 grant from the US National Institutes of Health it is 42 (refs. 2,3). Given the strong competition, what makes a good job candidate in academia? Roth opined that good publications, pedigree and recommendations suffice. But is that all it takes? Our faculty added the following. First, most Rockefeller faculty (as well as postdocs) surveyed felt that networking at meetings and seminars was important. We find this of note because David Roth, as well as most principal investigators (PIs) in a recent *Science Careers* survey[4], downplayed networking to an extent. Second, every single Rockefeller faculty member surveyed stated that good communication skills were crucial. This agrees with communication ranking first in a 2009 faculty survey on attributes of a successful postdoc[4,5]. Given that communication ranked number 7 in a national postdoc survey[4], postdocs should be aware of the importance faculty give to communication skills. More reason to keep using those lab meetings as practice ground! Teaching was another ingredient, but only for those applying to primarily teaching colleges. Mentoring undergraduates or summer interns, interestingly, did not make the cut.

We asked people to rank the relative weight of the following in choosing a job: location, salary, start-up package, department chair's leadership, and faculty profile. Roth, to our surprise, had brought up "the identity, vision and personality of the department chair" as his primary concern. Having a supportive, mentoring and visionary department chair was important to his decision, and is something he was glad to have considered. Some faculty thought that the start-up package and location of the university (to attract good students and postdocs) superseded the faculty profile and department chair. Postdocs, on the other hand, rated the faculty profile highest. Salary remained the bottom contender on everyone's list.

Competition arises from disparity between supply and demand. Does this imply that the system trains more people than science needs? People were divided on this question. Some thought that this may be affected by the lack of a retirement cutoff for 'baby boomer' faculty. Others suggested that if all careers after the PhD are included, we are training just the right amount.

Part of this issue is that many postdocs and faculty consider nonacademic positions as 'alternative careers' as opposed to a primary career path. It need not be so, but many of us struggle with the decision to not set up a lab. This decision is often fraught with much soul-searching and indecision, and given that seven out of ten postdocs are in this position, this is not trivial. Approximately 95% of postdocs felt that they would appreciate honest input from their PI on their ability to pursue an academic career. Every single PI polled, including Roth, said that they would suggest an alternative career to someone whom they felt was making a poor decision based on their strengths. We were surprised by this because we were not aware if this did really happen. Understandably, some faculty hinted that it may be unwise for them to pass personal judgment on someone's career—what if they are mistaken? We hope that this candid feedback will encourage honest discussions on career goals between postdocs and PIs. Finally, although many postdocs are keen to explore nonacademic careers, many feel unsure on how to go about this

and whom to tap for guidance. Be that as it may, most respondents said that finding a job in an alternative career remains the primary responsibility of the postdoc. PIs obviously cannot be our role model for a career outside the lab. The onus is on us, be it tapping into alumni networks or career networks such as the *Nature Network*, or proactively enhancing our resumes with skills we think we may need. What we can ask of and hope for from our PI is to be a sounding board, offering support and guidance in pointing us in the right direction.

Sustainability in running a lab. Two trends are gaining prominence in science: translational research and collaborative science. Are these just fads or a are they a shift in scientific methodology? Discussion participants agreed translational research is getting more hype than is warranted but does seem to be here to stay. We asked people to list the importance of translational research from "Primary goal of my research" to "basic research, without an attempt to translation, is just as good". Although some people did list it as critical, or somewhat important, it was interesting to note that the majority listed translational research as "somewhat important", ranging from "yes, important for NIH objectives" to "a way to fool granting bodies". Postdocs were much more likely than faculty to see translational research as important for publishing and funding.

Such science necessarily translates into collaborations—not all of us can with aplomb combine fields from genetics and biochemistry to tissue samples and clinical models. Not surprisingly, 85% of postdocs felt that multilab collaborations will increasingly become the norm, compared to 64% of faculty. The need for collaborations in the long run brings back to us the importance of networking, with our peers as much as with senior faculty.

We asked whether it was appropriate for junior PIs to collaborate with their previous mentors. The general response was that although it is not advisable, neither is it a strict 'no'. However, discussion participants indicated that in most cases, work done in collaboration with one's mentor is not likely to be taken as one's independent work—an important note for those pursuing tenure-track positions. Therefore, starting as postdocs, we should look for projects that will define our independent role and help augment our careers as postdocs today—and later our careers as independent investigators.

Sustainability of scientific temper. How do we know we will be able to sustain creativity, select and mentor students, and learn management of people and of costs? Is it a handicap to not have these skills honed during a postdoc? Although it may seem daunting, the faculty we surveyed said that, with common sense, developing these skills is easy, and lack of their prior development is not at all a handicap. However, they did highly recommend other skills as crucial: they suggested that postdocs ask to be involved in reviewing papers and writing grants. In addition, faculty suggested we should learn the skill of framing important questions with testable hypotheses, and stay focused on only the skills needed to get the job done. The bottom line: learn how to get grants; you will manage them just fine. Learn how to get papers out; new ideas will likewise follow just fine.

Summary

At the end of our discussion and survey, postdocs signed off wishing for more mentorship and interaction with their PI and more help when gearing up for nonacademic routes. We urge all faculty to be more involved in helping sketch out the career paths of their postdocs. At the same time, we also remind postdocs to become more proactive. If you miss some aspects of mentorship from your PI, it is wise to tap into alumni, peers or other PIs. It remains our responsibility to build sustainable careers—ones to last us a lifetime.

ACKNOWLEDGMENTS

The authors thank the postdocs and faculty at Rockefeller University for participating in their survey. We thank David Roth for discussions. The postdoc retreat was organized by the Rockefeller Univeristy Post-Doctoral Association.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. National Science Board. Science and Engineering Indicators 2006. <http://www.nsf.gov/statistics/seind06> (2006).
2. National Academy of Sciences. *Bridges to Independence* (National Academies Press, Washington, DC, 2005).
3. <http://grants.nih.gov/>.
4. Bonetta, L. The evolving postdoctoral experience. *Science Careers* (August 2009).
5. Bonetta, L. The postdoc experience. *Science Careers* (August 2008).

PEOPLE

VaxInnate (Cranbury, NJ, USA) has named **Thomas Hofstaetter** (far left) as president and CEO. **Alan Shaw** (near left), who previously held that position, has assumed the newly created role of CSO and becomes chairman of the board. Hofstaetter joined the company earlier this year as COO and member of the board after more than 30 years at research-based pharmaceutical companies. He previously served as senior vice president of corporate business development at Wyeth, where he also headed global business development for the company's pharmaceutical division. Shaw joined VaxInnate in 2005 from the vaccine division of Merck & Co., where he was executive director of the public policy, public health and medical affairs group.

"I am delighted to accept the challenge of taking VaxInnate and its paradigm-shifting vaccine technology platform to the next level," Hofstaetter said. "I look forward to collaborating with Alan and the entire VaxInnate team to accomplish that."

Pacific Biosciences (Menlo Park, CA, USA) has named **Susan K. Barnes** as CFO. Barnes has nearly 30 years experience in senior financial management. From 1997 to 2005, she was senior vice president, finance and CFO of Intuitive Surgical.

Craig W. Carlson has joined Hana Biosciences (S. San Francisco, CA, USA) as CFO. He has held senior leadership and financial management positions for the past 25 years, most recently as CFO and COO for 20 Cent Ventures. From 2006 to 2008, he served as CFO of Neurobiological Technologies.

Compugen (Tel Aviv) has announced the appointments of **Anat Cohen-Dayag** as president and CEO and **Martin Gerstel** as chairman. Cohen-Dayag and Gerstel previously served as co-CEOs of the company. Cohen-Dayag joined Compugen in 2002 and held the positions of director of diagnostics, vice president for diagnostic biomarkers and drug targets, and vice president of R&D. Gerstel joined Compugen in 1997 as chairman of the board, and served in that role until January of last year when he was appointed president and CEO, and then co-CEO in June. **Dov Hershberg**, who served as chairman of Compugen during 2009, will remain a director of the company.

Teva Pharmaceutical Industries (Jerusalem) has announced the appointment of **Phillip Frost** to serve as chairman of the company's board of directors, after the resignation of **Eli Hurvitz** to recover from illness. Frost previously served as vice chairman since 2006, after Teva's acquisition of Ivax, where he served as chairman and CEO. **Moshe Many**, who had been serving as interim chairman of Teva, has been appointed vice chairman of the board of directors.

Charles Homcy has announced his retirement as president and CEO of Portola Pharmaceuticals (S. San Francisco, CA, USA) to take on the role of co-chairman of the board of directors. He has served as president and CEO since he co-founded the company in 2003. COO **William Lis** will succeed Homcy as the new CEO and **Hollings C. Renton** has been appointed co-chairman and lead director of the board of directors. Renton previously served as president and CEO, director and chairman of the board of Onyx Pharmaceuticals.

Nicholas A. LaBella Jr. has joined Insmed (Richmond, VA, USA) as CSO. LaBella has held a number of senior-level positions in drug development and regulatory affairs at Watson Laboratories and the former Sandoz Research Institute.

OXiGENE (S. San Francisco, CA, USA) has announced the resignation of **Arthur B. Laffer** from its board of directors to pursue other longstanding interests. Laffer has served on the OXiGENE board since 1998 and had served as chair of the audit committee, the chair of the compensation committee and a member of the nominating and governance committee.

Rib-X Pharmaceuticals (New Haven, CT, USA) has announced the appointment of **Mark Leuchtenberger** (left) as president, CEO and a member of the board of directors. He joins the company from Targanta Therapeutics, where he served as president and CEO. His predecessor as CEO, **Susan Froshauer**, has been appointed to the newly created role of CSO at Rib-X.

Prosensa (Leiden, The Netherlands) has announced the appointment of **Berndt Modig** (left) as CFO. He brings more than 20 years of experience in company finance to Prosensa, most recently serving as CFO of Jerini.

Adeona Pharmaceuticals (Ann Arbor, MI, USA) has named **Jeff Lucero Riley** as a member of its board of directors. Riley has more than 19 years of experience in the biotech and pharma industries. He is currently managing director of Black Crow Ventures and has held senior corporate and commercial development positions at Amphora Discovery, Ontogen and AvMax.

Cytokinetics (S. San Francisco, CA, USA) has announced the resignation of co-founder and former CEO **James H. Sabry** from its board of directors; Sabry is taking an operational role at a pharmaceutical company. Sabry, who had been serving as chairman of the board, will remain a consultant to the company. **L. Patrick Gage**, formerly president of Wyeth Research and a Cytokinetics board member since 2009, has been named to succeed Sabry as chairman.