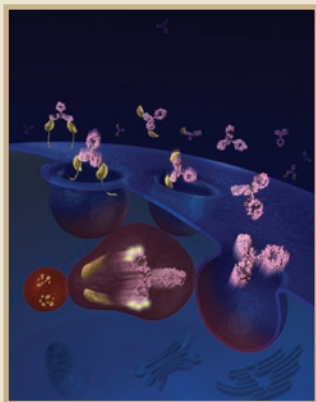


nature biotechnology



Therapeutic IgG antibodies bind only two antigen molecules during their lifetime. Igawa *et al.* engineer antibody recycling through pH-dependent binding, which allows the antigen to be released in the endosome for lysosomal degradation while the antibody is returned to the plasma for additional rounds of antigen binding (p 1203). Credit: JOLLYBOY and Sakura Motion Picture Co., Ltd.

© 2010 Nature America, Inc. All rights reserved.



Coverage with evidence development, pp1157 and 1160



nature publishing group

EDITORIAL

1133 **Conflicts and collaborations**

NEWS

- 1135 **Novartis eyes oral MS drug as potential blockbuster**
- 1137 **Bristol-Myers Squibb reaps biologics in ZymoGenetics windfall**
- 1139 **Shire's replacement enzymes validate gene activation**
- 1139 **Pharmacogenomics row**
- 1140 **Joint inspections still cool**
- 1140 **Stimulus trickle**
- 1141 **Transgenic salmon inches toward finish line**
- 1141 **CIRM spurs translation**
- 1141 **Irish bait**
- 1142 **Sanofi/Genzyme hostile**
- 1142 **Adverse-events fraud trial**
- 1143 **NEWSMAKER: Anaphore**
- 1144 **DATA PAGE: Biotech rallies in Q3**
- 1145 **NEWS FEATURE: When patients march in**

BIOENTREPRENEUR

BUILDING A BUSINESS

- 1149 **Making the leap into entrepreneurship**
Randall Schatzman, Mark Litton, John Latham & Jeffrey Smith

OPINION AND COMMENT

CORRESPONDENCE

- 1152 **Grant management skills are critical for young scientists**
- 1153 **Chinese hamster ovary cells can produce galactose- α -1,3-galactose antigens on proteins**

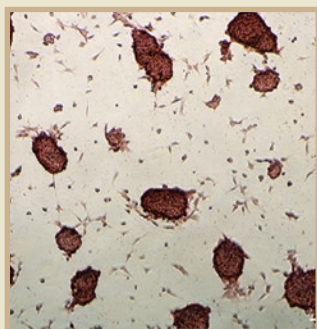
COMMENTARY

- 1157 **A policy approach to the development of molecular diagnostic tests**
Kevin A Schulman & Sean R Tunis
- 1160 **What is the value of oncology medicines?**
Joshua Cohen & William Looney

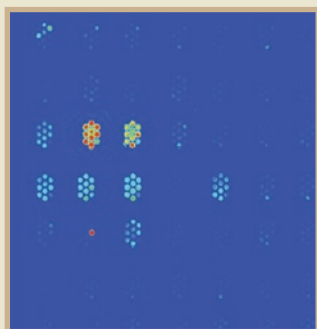
Nature Biotechnology (ISSN 1087-0156) is published monthly by Nature Publishing Group, a trading name of Nature America Inc. located at 75 Varick Street, Fl 9, New York, NY 10013-1917. Periodicals postage paid at New York, NY and additional mailing post offices. **Editorial Office:** 75 Varick Street, Fl 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. Copyright © 2010 Nature America, Inc. All rights reserved. Printed in USA.



Culturing innately undifferentiated plant cells, p 1175



Chondrocyte-like cells from ES cells, p 1187



Antibodies discovered from small libraries, p 1195

FEATURE

1165 What's fueling the biotech engine—2009–2010

Saurabh Aggarwal

PATENTS

1172 A shadow falls over gene patents in the United States and Europe

Gareth Morgan & Lisa A Haile

1174 Recent patent applications in biological imaging

NEWS AND VIEWS

1175 Plant natural products from cultured multipotent cells

Susan Roberts & Martin Kolewe **b** see also p 1213

1176 Making antibodies from scratch

J Christopher Love **b** see also p 1195

1178 Toward global RNA structure analysis

David M Mauger & Kevin M Weeks

1180 RESEARCH HIGHLIGHTS

COMPUTATIONAL BIOLOGY

COMMENTARY

1181 *In silico* research in the era of cloud computing

Joel T Dudley & Atul J Butte

RESEARCH

ARTICLES

1187 Directed differentiation of human embryonic stem cells toward chondrocytes

Rachel A Oldershaw, Melissa A Baxter, Emma T Lowe, Nicola Bates, Lisa M Grady, Francesca Soncin, Daniel R Brison, Timothy E Hardingham & Susan J Kimber

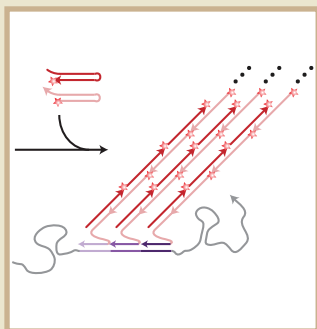
1195 Spatially addressed combinatorial protein libraries for recombinant antibody discovery and optimization

Hongyuan Mao, James J Graziano, Tyson M A Chase, Cornelia A Bentley, Omar A Bazirgan, Neil P Reddy, Byeong Doo Song & Vaughn V Smider **b** see also p 1176

LETTERS

1203 Antibody recycling by engineered pH-dependent antigen binding improves the duration of antigen neutralization

Tomoyuki Igawa, Shinya Ishii, Tatsuhiko Tachibana, Atsuhiko Maeda, Yoshinobu Higuchi, Shin Shimaoka, Chifumi Moriyama, Tomoyuki Watanabe, Ryoko Takubo, Yoshiaki Doi, Tetsuya Wakabayashi, Akira Hayasaka, Shoujiro Kadono, Takuya Miyazaki, Kenta Haraya, Yasuo Sekimori, Tetsuo Kojima, Yoshiaki Nabuchi, Yoshinori Aso, Yoshiki Kawabe & Kunihiro Hattori



Multiplexed *in situ* hybridization,
p 1208

- 1208 Programmable *in situ* amplification for multiplexed imaging of mRNA expression**
Harry M T Choi, Joann Y Chang, Le A Trinh, Jennifer E Padilla, Scott E Fraser & Niles A Pierce
- 1213 Cultured cambial meristematic cells as a source of plant natural products**
Eun-Kyong Lee, Young-Woo Jin, Joong Hyun Park, Young Mi Yoo, Sun Mi Hong, Rabia Amir, Zejun Yan, Eunjung Kwon, Alistair Elfick, Simon Tomlinson, Florian Halbritter, Thomas Waibel, Byung-Wook Yun & Gary J Loake
b see also p 1175

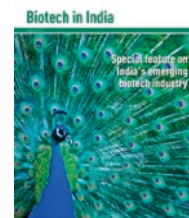
CAREERS AND RECRUITMENT

- 1218 Third quarter biotech job picture**
Michael Francisco
- 1220 PEOPLE**

ADVERTISEMENTS

Special Report: Biotech in India

India is known to be very strongly positioned in the manufacture of generics. Is it now ready to move up the value chain as an innovator? This supplement looks at the rapid progress in the Indian biotech market and its positioning as an innovation partner to the global industry. This special report is produced with the commercial support of the Department of Biotechnology (DBT); the Association of Biotechnology Led Enterprises (ABLE); MMAActiv, a Sci-Tech Communication Company, and the other advertising organizations featured. The special report follows Letters after page 1217 and is produced with commercial support from the promoting organizations within the section.



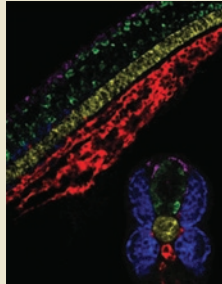
Special Report: BioPharma Dealmakers— Not-for-Profit, Good for Business

Can partnering with not-for-profits really be good for business? Freelance writer George Mack and supplement editor Barbara Nasto investigate by looking at deals between commercial organizations and not-for-profits. What they uncover is that partnering with not-for-profits is good for business on many levels. The report follows Special Report: Biotech in India and is produced with commercial support from the promoting organizations within the section.



Multiplexed *in situ* hybridization

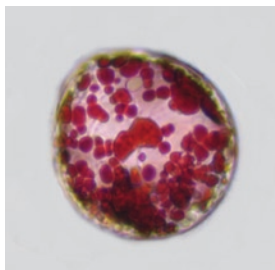
The dissection of the complex regulatory processes that control the spatial organization of multicellular organisms relies on the analysis of expression patterns of genes *in situ*. A major limitation of current mRNA imaging technologies is their limited ability to visualize more than one or two mRNAs simultaneously in complex tissues. Pierce and colleagues present a method that detects mRNAs using polymerization of fluorophore-labeled RNA hairpins in a hybridization chain reaction. The polymerization is triggered by sequence-specific RNA probes fused to an initiator sequence that also tethers the growing polymer to the mRNA target. The authors show that at least five different RNAs can be visualized at the same time in whole-mount embryos with orthogonal polymerization reactions. They achieve deep penetration of their probes and high signal-to-background ratios. Potential applications of the method include analysis of signaling pathways during development and pathogenesis, characterization of rare cell types and diagnostic staining of tissue samples. [Letters, p. 1208]



ME

Natural products from cultured multipotent plant cells

The plant kingdom offers a plentiful supply of secondary metabolites, many of which have been shown to be of value as drugs, pigments, fragrances and pesticides. However, natural products are often produced only in trace amounts or are sometimes found only in endangered or slowly growing species. Because the complexity in the structure of plant natural products often defies their production by synthetic chemistry or metabolic engineering of microbes, the most common approaches to produce valuable plant natural products on an industrial scale have involved suspension-cultured plant cells. These lines, which are generally thought to be derived from a dedifferentiation process, consist of a heterogeneous mixture of cell types that fail to sustain acceptable rates of growth and natural product synthesis with repeated subculture. Working with four species, Loake and colleagues report that cultures comprising only meristematic cells from the



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

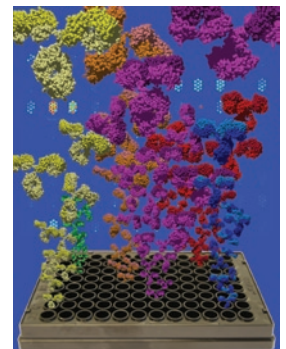
cambium, the zone of multipotent cells that give rise to the plant vascular tissues xylem and phloem, may provide a solution to these challenges for at least certain natural products. Having first obtained compelling evidence that cambial meristematic cells (CMCs) from *Taxus cuspidata* have stem cell-like properties, they demonstrate that the CMCs may provide a cost-effective strategy for commercial production of the anti-cancer agent paclitaxel. Similarly promising results are shown for the production of ginsenoside F2 from CMCs derived from *Panax ginseng*. [Letters, p. 1213; News and Views, p. 1175]

PH

One-by-one antibody screening

Antibodies that modulate the activity of a target are difficult to discover with display-based approaches, which select only high-affinity binders. Smider and colleagues identify antibodies with a range of affinities using a small-molecule discovery method that involves one-by-one screening of an optimized small library of antibody fragments with known sequences. The library was constructed by first designing *in silico* a set of heavy and light chain genes that was as diverse as possible. Then DNA sequences encoding the genes were synthesized, cloned into plasmids and expressed in *Escherichia coli*, and purified protein was recovered. About 10,000 antibody fragments were screened against 9 antigens, identifying 85 hits against 7 antigens. Notably, the one-by-one screening approach enabled the authors to compare the binding affinities of hits with those of antibodies having similar sequences to identify residues to target for further optimization. This work argues against the prevailing view that billion-member libraries are required to discover new antibodies. [Articles, p. 1195; News and Views, p. 1176]

CM



Recycle, reuse, reduce

For many therapeutic antibodies, a high rate of antigen-mediated clearance necessitates the use of relatively high doses and frequent drug administration. Igawa and colleagues engineer tocilizumab (Actemra), a humanized antibody against the IL-6 receptor (IL-6R), to dissociate from IL-6R within the acidic environment of the endosome without compromising its affinity for IL-6R in plasma. The rationale behind this strategy is that after dissociation within the endosome, IL-6R is targeted for lysosomal degradation and the antibody is recycled back to the plasma for additional rounds of antigen neutralization. Consistent with this notion, tocilizumab engineered for pH-dependent antigen-binding displays improved pharmacokinetics and IL-6R neutralization in both transgenic mice and cynomolgus monkeys. Engineering 'histidine switches' may improve the clinical potential of other antibodies used to treat chronic diseases, especially when used in conjunction with affinity maturation and other strategies to extend *in vivo* longevity. [Letters, p. 1203]

PH

Chondrocytes from ES cells

Research on tissue engineering to repair articular cartilage has mostly ignored embryonic stem (ES) cells because of the difficulty of converting them to chondrocytes in high numbers. Kimber and colleagues present a method for differentiating human ES cells toward

chondrocytes that is efficient and chemically defined, avoiding feeder cells and serum. The authors look to developmental biology to design a step-wise approach in which the ES cells transition through primitive streak–mesendoderm and mesoderm on the way to chondrocyte-like cells. At the end of the process, the majority of the cells express the chondrocyte marker *SOX9*, and there is abundant production of cartilage components such as sulfated glycosaminoglycan, collagen type II and cell-surface CD44. With further optimization, the approach may be valuable in the development of cartilage regeneration strategies. [Articles, p. 1187] KA

Patent roundup

A hotly contested platform for generating human proteins—the subject of lawsuits filed by Amgen and Genzyme against the technology's developer TKT—has finally found its way to market with two enzyme replacement therapy approvals for Shire, which bought the platform. [News, p. 1139] LM

Will recent rulings that restrict DNA-based patents reduce financial incentives and chill investment? Morgan and Haile write that in Europe, *Monsanto v. Cefetra* has legal practitioners trying to work around a law that unfairly discriminates against DNA-based inventions. Meanwhile, in the United States, the industry awaits the next step in *Association for Molecular Pathology, et al. v. USPTO*. [Patent Article, p. 1172] MF

Recent patent applications in biological imaging. [New patents, p. 1174] MF

Next month in

nature biotechnology

- Nanoparticle behavior after lung delivery
- Coalescence of DNA reading and writing
- Simpler preparation of minicircles
- Improved gene synthesis from cheap oligos
- Sterile insects suppress resistance to *Bt* cotton



Conflicts and collaborations

Claims of conflicts of interest concerning authorship of a scientific paper highlight the difficulties facing regulators participating in collaborations with industry.

Back in mid-August, the US Department of Health and Human Services (DHHS) stated that it was investigating a conflict-of-interest allegation involving Janet Woodcock, the director of the US Food and Drug Administration's (FDA) Center for Drug Evaluation and Research (CDER). The investigation involves a complaint filed by Amphastar Pharmaceuticals that one of its competitors, Momenta Pharmaceuticals, had privileged access to Woodcock during the approval process for a generic version of Lovenox, a low-molecular-weight heparin (LMWH) product.

Amphastar claims to be hard-done-by because it was the first of the two companies to submit its abbreviated new drug application for enoxaparin sodium way back in March 2003. And although Amphastar's drug is yet to be approved, Momenta received approval for its product in July, even though it filed two years later in August 2005. In the meantime, Janet Woodcock and scientists at Momenta were co-authors on a scientific paper published in a high impact journal—an interaction that Amphastar lawyers have raised as one of the key pieces of evidence of the inappropriate closeness between FDA staffers and Momenta personnel.

At this point, we must also declare our interests. The journal in which the paper appeared was *Nature Biotechnology* (26, 669–675, 2008). The research dealt with analytical methods for assessing the potential presence of highly sulfated chondroitin compounds in preparations of heparin, contaminants that were responsible for 150 deaths in patients who experienced an allergic or hypotensive response after being exposed to the drug. The heparin in question was in products from Baxter produced using active ingredients sourced from a company called Scientific Proteins Limited (SPL), in China. Two of the authors on the paper, Robert Langer and Ram Sasisekharan, were not only directors of, and stockholders in, Momenta, but also advisors to SPL at the time of publication.

One of the realities in this affair is that, as a result of the adverse events arising from contaminated heparin, the approval applications of both companies—Momenta and Amphastar—were held up. Without access to the correspondence with FDA, little can be said about the reasons for the different timelines on the applications—further information may come to light following completion of the DHHS investigation that is still underway. However, on the broader question of conflicts arising from industry-regulator collaborations, much more can be said.

First, industry-regulator co-publication is not a new phenomenon. Although, until the mid-1990s, FDA-industry co-publications were quite rare—just three a year or fewer, from 1996 onwards—the number has steadily risen so that since 2003, between 20 and 49 co-publications have appeared per year. Many, but by no means all, are 'methods' papers, describing, for instance, the development of biochemical and cellular assays or the use of new animal models.

Although the growth in co-authored papers is hard to pin down to one factor, it seems likely to be associated with the passage of PDUFA

legislation in 1992, which allowed FDA to levy user fees from industry. The association is certainly subtle: PDUFA fees allowed FDA to increase staffing levels, to become more efficient at processing applications and to add industry-standard 'competences'. Thus, a PDUFA-powered FDA has been able to participate in peer-to-peer discussion on matters such as mechanisms of action, potential side effects and the use of new categories of data applied to new types of interventions. And, in the interests of spreading their knowledge, FDA has published this work in the scientific literature. Indeed, FDA publishes a lot of journal articles—over 15,000 in the past 15 years, only 2% of which had industry co-authors.

These peer-to-peer discussions have entered into a new dimension with the arrival of R&D programs, such as the Critical Path Initiative in the United States or the Innovative Medicines Initiative (IMI) in Europe. Critical Path is a joint research program with an agenda set jointly by industry and the FDA and designed to anticipate some of the technical issues that hold back drug development. IMI is funded by the European Commission and (in kind) by a consortium of pharma companies. Industry sets the IMI agenda, but regulatory bodies participate in many of the consortia that IMI funds. Although the schemes are structured differently, both serve to bring industry and the regulators closer together.

The major issue—whether industry research collaborations increase the likelihood of conflicts of interests for agency staffers—has received relatively little attention in the wider media. Even so, the FDA has made efforts to compartmentalize responsibilities for regulatory oversight away from collaborative research. Thus, in the Predictive Safety Testing Consortium (PSTC), which is led by Critical Path, the regulators involved in assessing the scientific findings of the collaboration were different from those who worked with industry researchers on the experiments. And a different set of staffers again are involved in overseeing drug filings from the companies that are partners with FDA in PSTC. Such 'Chinese walls' guard against conflicts.

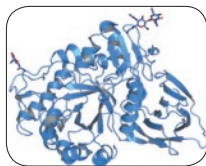
In the case of Janet Woodcock and Momenta, Amphastar went so far as to hire the private detective agency Kroll to collect details about her family, retrace the steps she made on business trips and file Freedom of Information Act requests for Woodcock's e-mails, phone records, voice-mails, calendar and expense reports. These actions step over the line, but the FDA's response has also been less than satisfactory. Earlier in the year, FDA legal counsel Ralph Tyler announced "We've determined that there's no conflict here" and stated that Woodcock had stepped aside in 2009 from any involvement in the LMWH applications.

But the fact is, it was inappropriate for Woodcock to author a paper with a company filing products for review at CDER, where she is the director. The unfortunate reality is that these days, even the perception that a conflict might exist is sufficient to damage the reputation of an agency that already has fallen far in the public's estimation. **LB**

IN this section



**Bristol-Meyers
Squibb snaps up
ZymoGenetics**
p 1137



**Shire's alternative
to recombination
succeeds** p 1139



**Agency weighs up
transgenic salmon
approval** p 1141

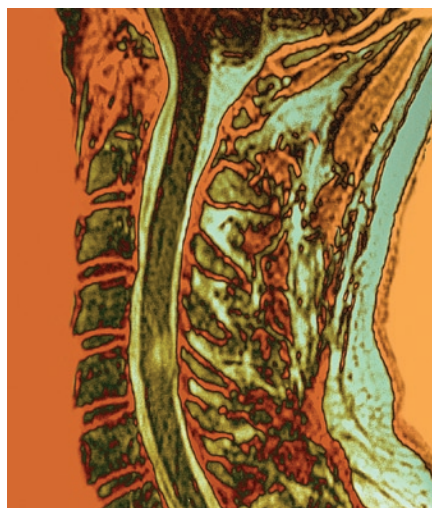
Novartis eyes oral MS drug as potential blockbuster

The first oral medication to treat multiple sclerosis (MS) has been given a regulatory green light after getting an overwhelmingly positive recommendation by an advisory panel three months earlier. In September, the US Food & Drug Administration (FDA) approved Novartis's Gilenya (fingolimod, also called FTY720) as a first-line treatment for people with the relapsing form of MS. The decision follows extensive discussion over dosing and safety issues related to the drug's novel mechanism, which sequesters lymphocytes in the lymph nodes to prevent the inflammation and tissue damage that characterize MS. Basel-based Novartis wasted no time launching the small-molecule drug, which it has been touting for some time as a potential blockbuster.

The newly approved drug provides "a compelling treatment option to needle-phobic patients," says Joel Sendek of Lazard Capital Markets in New York. Over time, it will take share away from other treatments, he says, "as patients discouraged with monthly infusions elect to switch, and physicians and patients gain comfort with the safety and tolerability profile."

Gilenya—approved to reduce relapses and delay disability progression—is competing with Tysabri (natalizumab), which Cambridge, Massachusetts firm Biogen Idec co-markets with Dublin-based Elan, as well as several formulations of interferon (IFN)- β 1a, for a share of the \$11 billion, and growing, MS market. There are 400,000 people in the US with MS and over 2 million worldwide. The US go-ahead follows the drug's approval in Russia earlier in the month; it is still under consideration by European regulators.

Analysts anticipate Gilenya will snatch market share from other drugs—in particular, Biogen's Tysabri. Sendek estimates Gilenya sales of \$392 million in 2011 and \$798 million in 2012. A physicians' survey, conducted by Lazard, estimates that preference for an oral agent will sharply reduce Tysabri's use as a second-line treatment. A June survey by RBC Capital Markets in San Francisco, just after the Gilenya FDA panel meeting, came to a similar conclusion. The data suggested that 75% of neurologists would opt for Novartis's drug as first choice over Tysabri, which is used primarily



A spinal MRI scan through the neck and upper back of a person with multiple sclerosis. A slew of compounds in late stage development will broaden therapeutic options for this indication in the next few years.

only when other treatments have failed. "We expect a significant drop in net new patients on Tysabri," says RBC analyst Jason Kantor, starting in the fourth quarter of 2010.

Kantor also says that Gilenya will displace Tysabri because physicians sometimes impose a 'treatment holiday' when using that drug, a practice that is becoming more common due to the risk of progressive multifocal leukoencephalopathy, which has dogged Tysabri for years (*Nat. Biotechnol.* 27, 986, 2009). According to the RBC survey, he says, "approximately half would rather switch to Gilenya than restart Tysabri."

Gilenya, though, is not without its own safety concerns. The FDA has required an additional post-approval surveillance of the drug to gauge effectiveness at a dose lower than the one approved and also has stipulated a five-year, worldwide safety study to be run. What's more, physicians will need to monitor patients after the first dose for possible cardiovascular side effects.

As a result, it may take some time to get physicians to switch to Gilenya when other treatment options exist. In trials, the drug

has shown side effects that include a slowed heart rate, macular edema and breathing and liver problems. Although these may be addressed by monitoring, physicians may choose to continue with current treatments as they await results of Novartis's post-approval safety study. Until then, Gilenya will likely be used as an option mostly for newly diagnosed MS or in individuals who are not doing well with current medications.

"There's still a concern that by holding cells in the lymph nodes, you may be opening up patients to other situations such as skin cancer and maybe some infection," says advisory panel member Jacqueline Friedman of the New York University Langone Medical Center. "Although having gone to the [panel] meeting, it doesn't seem like a huge risk, it is an unknown box."

What became clear very quickly at the FDA advisory panel meeting is that the drug's effectiveness is not at issue. In one previously reported phase 3 trial, Gilenya reduced the rate of relapse by 52% compared with IFN- β 1a, with significantly greater reductions in brain atrophy as well. In another published phase 3 trial at the 0.5 mg dose that FDA approved, the drug reduced the risk of disability progression by 30% over placebo at three months and 37% at six months and also significantly reduced the rate of brain atrophy over the two-year study period.

On the basis largely of these data, the panel concluded by a vote of 25–0 that Gilenya had demonstrated effectiveness to reduce the frequency of clinical exacerbations of MS, and also, by a vote of 24–1, that it was effective in delaying the onset of physical disability. (Friedman, the one who dissented, did so only on the basis that one-year data was not sufficient to draw the conclusion.)

"Standing alone, the results indicate it's a highly effective therapy," says Fred Lublin of New York's Mt. Sinai Medical Center. Lublin, a consultant to Novartis, presented the clinician's perspective on Gilenya at the panel meeting.

But concerns over the potential for viral infections do linger, even if they don't seem as worrisome as they did two years ago, when one patient died from varicella zoster (chicken

Table 1 Selected multiple sclerosis products in development

Company	Drug	Description	Target	Status
Genzyme	Campath (alemtuzumab)	Humanized monoclonal antibody (mAb)	CD52 on leukemia cells	Phase 3
Biogen IDEC	Daclizumab	Humanized mAb	Alpha subunit of Interleukin-2 receptor of T cells	Phase 3
Roche (Basel)	Ocrelizumab	Fully humanized mAb	CD20 on mature B cells	Phase 3
Biogen IDEC	BIIB017	PEGylated interferon β 1a	Interferon receptor	Phase 3
Actelion (Allschwil, Switzerland)	ACT-128800	Small-molecule receptor agonist	Sphingosine-1 phosphate receptor (S1P-R)	Phase 2b
Bayhill (Palo Alto, California)	BHT-3009	A tolerizing DNA vaccine	Myelin basic protein	Phase 2b
Novartis	AIN457	Neutralizing IgG1K mAb	Interleukin 17	Phase 2
GlaxoSmithKline (London)	Arzerra	Fully human, high-affinity antibody	CD20 molecule on B cells	Phase 2
GlaxoSmithKline	Firategrast	Small-molecule antagonist	Integrin α -4 β -1/VLA-4	Phase 2
Teva (Petach Tikva, Israel)	ATL/TV1102	Antisense, inhibits cellular production of adhesion molecule VLA-4	Integrin α -4 β -1/VLA-4	Phase 2
Novartis	BAF312	Small-molecule immunomodulator	S1P-R	Phase 2
Eli Lilly (Indianapolis)	LY2127399	Fully human neutralizing mAb anti-B cell-activating factor (BAFF)	BlyS/BAFF/TACI/BCMA receptor	Phase 2
Ono Pharmaceutical (Osaka, Japan)	ONO-4641	Small-molecule immunomodulator	S1P-R	Phase 2

Source: Sagient Research/BioMedTracker

pox) and another went into coma due to herpes encephalitis. These incidents led to increased monitoring for infections (*Nat. Biotechnol.* **26**, 844–845, 2008).

At the panel meeting, Judith Feinberg of the University of Cincinnati said she was “a little flustered by the fact that the primary mechanism of action was not very much investigated from the standpoint of toxicity to patients.... I’m trying to understand how we have a drug whose primary mechanism doesn’t permit egress of T cells from the thymus or T and B cells from the lymph nodes and we have such a complete absence of data of what happens specifically to B cells and T-cell subsets in these patients.”

In MS, lymphocytes cross the blood-brain barrier and attack the central nervous system. The process leads to inflammation, nerve damage and loss of function from a combination of demyelination of the protective sheath around nerve cells, loss of axons and scarring. Although other MS drugs eliminate lymphocytes, Gilenya traps them in the nodes. Gilenya does so by binding sphingosine-1 phosphate (S1P) receptors. S1P is a naturally occurring lipid on cells of the immune, cardiovascular and central nervous systems, which regulates the exit of lymphocytes into circulation.

Of all five S1P subtypes, the most prominent on lymphocytes is S1P-1. “That’s the subtype that results in the trapping of B cells,” says Trevor Mundel, head of worldwide clinical development at Novartis. Binding S1P-5, which is expressed in neurons and glial cells, is also helpful in MS. “Preclinically at least, it is related to some factors around growth and survival of B cells,” says Mundel, “which is one of the reasons why we felt that drugs like Gilenya probably have CNS [central nervous system] effects as well, which we established in animal models and *in vitro* systems.”

Gilenya’s unwanted cardiovascular side effects are related to its targeting of S1P-3, which is found in the heart and smooth muscle.

Another issue for the regulators was that Gilenya, which first entered clinical trials as an immunosuppressant in transplantation—at much higher doses—had not been tested at a dose lower than the now-approved 0.5 mg. As part of the approval, FDA told Novartis to conduct a trial at 0.25 mg, to see if the drug is effective at that dose, and whether the safety profile changes. “I always wanted to do that trial,” says Mundel. “It was always on the books.”

The same week as Gilenya’s approval in the US, Europe’s Committee for Medicinal Products for Human Use recommended against approving Movectro (cladribine), an oral lymphocyte-depleting purine nucleoside analog from EMD Serono of Darmstadt, Germany, indicated for MS. The agency cited the risk of lymphopenia and several cases of cancer that arose during

trials. The decision was somewhat surprising given the drug’s approval in Australia earlier in the month, as well as its approval in Russia in July 2010. It gives Novartis more room to corner the oral MS drug market.

In addition to competition from Movectro, Novartis will be keeping an eye on Cambridge, Massachusetts-based Genzyme, which is testing its cancer antibody Campath (alemtuzumab) in MS (Table 1). Campath binds CD52 on the surface of lymphocytes and Genzyme has been pointing to Campath’s potential in MS as one reason why Sanofi-aventis’s proposed takeover bid for the company is too low (this issue, p. 1142).

But safety concerns will linger with any of these drugs. “If we limit the release of lymphocytes into the system generally, we may be opening the patients up to other infections or cancers,” says Friedman.

Mark Ratner Cambridge, Massachusetts

IN their words



“We will have all the pieces ready to go, and on the day we will just push the button.”

Novartis head of vaccine research Rino Rappuoli comments on how a collaboration with Craig Venter’s Synthetic Genomics Vaccines aims to slash the time

needed to start producing influenza vaccines. (*Bloomberg*, 7 October 2010)

“Seems like we’re always looking over our shoulder to see what’s next.” Jack Mosher, assistant research scientist at the University of Michigan, sums up the anxiety hanging over

the field of embryonic stem cell research as opponents attempt to ban federal funding. (*The Detroit News*, 5 October 2010)

“All of a sudden I’ve taken a five-year learning curve and shortened it to six months,” Laura Esserman, at the University of California, San Francisco, explains how in the I-SPY-2 breast cancer trial, which she co-leads, results are analyzed right away, unlike conventional trials where no one sees the data until the end. (*Wall Street Journal*, 2 October 2010)

“This is not the biggest thing on their plate. But in terms of the country itself, this is probably one of the biggest things we’ve got confronting us right now.” Tom Harkin, Iowa Senator and HHS appropriations chair, is frustrated by the FDA’s lack of resources to push for cell-based flu vaccines. (*The RPM Report*, 11 October 2010)

Bristol-Myers Squibb reaps biologics in ZymoGenetics windfall

Following a \$885 million buyout announced on 7 October, Bristol-Myers Squibb (BMS) of New York now has added Seattle-based ZymoGenetics to its biologics portfolio. Some critics have chided ZymoGenetics' board of directors for selling the company too cheaply to the pharmaceutical giant, which last year also acquired antibody pioneer Medarex and its rich pipeline of antibody products. Indeed, although ZymoGenetics lacked marketing muscle for its single approved product, Recothrom (recombinant thrombin), it had at least two products in development that looked very promising. Just three weeks after the buyout was announced, ZymoGenetics released impressive results for its interleukin-21 (IL-21) for metastatic melanoma. And positive early clinical results for a novel interferon (IFN)- λ for the treatment of hepatitis C virus (HCV) have raised the question, Why didn't ZymoGenetics management hold off from a deal until completion of the phase 2 trial of IFN- λ ?

"Perhaps \$885 million now would be worth more than a billion dollars at some indefinite time," says Allan B. Haberman, principal of Haberman Associates of Wayland, Massachusetts, referring to the figures the biotech stood to gain through its existing agreement with BMS signed in January 2009 to co-develop IFN- λ decorated with polyethylene glycol (PEGylated). Those milestones,

however, would have depended on the success and timing of the results of phase 2 and 3 trials and registration with the US Food and Drug Administration (FDA).

The main impetus for the buyout, most analysts believe, was to gain full control of ZymoGenetics' PEGylated-IFN- λ . With four antivirals in development and four on the market, BMS already enjoys a stronghold in the HCV market but one that may soon be challenged (**Box 1**). "Despite a very strong antiviral marketed portfolio, pressures on its marketed drugs means a targeted acquisition is what BMS needed," says Aparna Krishnan, analyst at IHS Global Insight in Lexington, Massachusetts.

But ZymoGenetics' pipeline riches extend beyond PEGylated-IFN- λ . The biotech's melanoma candidate denenicokin (recombinant IL-21, a cytokine that regulates cytotoxic T cells and natural killer cells), currently in phase 2b, is another important component of the deal, says Joshua Ovide, pharmaceutical analyst at Datamonitor, London. On September 29, ZymoGenetics reported the results of an open label phase 2a trial with recombinant IL-21 in patients with stage IV melanoma. The treatment led to a median overall survival of 12.4 months, with 53% of the 39 patients alive at 12 months.

In addition to a promising melanoma treatment, BMS also brings on board six early-stage biologics outlicensed to either Merck Serono



Eben Calhoun

ZymoGenetics headquarters building viewed from Lake Union, Seattle.

Box 1 Small molecules shake HCV market

As the race to secure a bigger tranche of the growing hepatitis C virus (HCV) market intensifies, it is not surprising that BMS's interest in ZymoGenetics hinged on securing sole ownership of PEGylated-IFN- λ . The HCV market, according to market research firm Global Data, was worth \$4 billion in 2009 and is projected to reach \$8.5 billion by 2016. Current standard of care for hepatitis C patients is a combination of the immune modifier PEGylated-IFN- α and ribavirin (Rebetol, Copegus, Virazole), an antiviral nucleoside analog that synergizes the IFN's effects. But the treatment not only has limited efficacy in individuals infected with genotype 1 HCV—the biggest market and unmet need—but also typically involves 48 weekly injections that can cause numerous side effects (*Nat. Biotechnol.* 27, 964, 2009).

Expectations are running high for novel IFN products, such as PEGylated-IFN- λ . The hope is that this product, which mediates activity through IL-28Ra (an IFN receptor different from that of IFN- α), will deliver the same treatment responses with fewer side effects. Interim phase 1b results for treating patients with PEGylated-IFN- λ and ribavirin were positive, but it is still early days for the drug.

Meanwhile, the large HCV patient population continues to attract other companies into the arena. Around 29 specifically targeted candidates are in phase 2 and 3 trials, including small-molecule inhibitors of the viral nonstructural serine protease 3 (NS3) and NS5 polymerase. Ahead of the game are Vertex of Cambridge, Massachusetts, with the small-molecule

antiviral telaprevir (VX-950) and Merck of Whitehouse Station, New Jersey, with boceprevir (SCH503034), both in phase 3 trials. HCV-NS3 protease inhibitors are expected to file at the end of 2010; telaprevir is thought to have the edge because it can dramatically increase the cure rate while also halving treatment time, says Mansi Shah, senior healthcare analyst at Datamonitor. Phase 3 results from Vertex released in September show that, when used in conjunction with the standard regime, telaprevir clinically cured 65% of patients who had previously been treated unsuccessfully, compared with only 17% for the standard regimen.

Despite their strengths, telaprevir and boceprevir are likely to be approved for use in combination with ribavirin and PEGylated IFN, products that are expensive, inconvenient and have numerous side effects. As a result, “the market is moving towards eliminating the PEGylated-IFN backbone of treatment regimens,” explains Shah, who says companies are looking to combine products so that regimens are IFN-free. Gilead of Foster City, California, and Vertex, for example, are both testing cocktails of protease and polymerase inhibitors in phase 2 trials.

Although such combinations could limit the commercial opportunity for PEG-IFN- λ , BMS could still be in the running as it has three other hepatitis C candidates in early clinical development and is also testing BMS-790052, its NS5A inhibitor, in combination with its protease inhibitor BMS-650032 in phase 2 trials. **ED**

of Geneva, or Novo Nordisk of Bagsvaerd, Denmark. These include ataccept—a fusion of the modified Fc domain (hinge-CH2-CH3) of human IgG1 to the extracellular region of TAC1 (transmembrane activator of the tumor necrosis factor family, calcium modulator and cyclophilin ligand interactor), which binds B-lymphocyte stimulator—which is in phase 3 development for systemic lupus erythematosus, and an anti-IL-31 humanized monoclonal antibody (mAb) for psoriasis not yet in the clinic.

“The biggest gain by far for BMS is in the drug discovery pipeline,” says Krishnan. “The addition of six novel biologics in various stages of clinical development boosts the commercial prospects for [BMS] in the medium term.” The partnered programs will bring in milestone payments and royalties; collaborations and licensing accounted for 70% of the \$73.22 million that ZymoGenetics registered in revenues for the first half of 2010, according to Krishnan. ZymoGenetics also markets Recothrom (a recombinant human thrombin to treat surgical bleeding) and has seven outlicensed commercial products, including Novolin (insulin) and NovoSeven (factor VIIa).

BMS is expected to feel the full extent of generic competition in 2012, when (among others) its patent on the blood thinner Plavix (clopidogrel) is expiring. To offset this, the company has divested all of its non-pharma activities (most recently its Mead Johnson nutritionals business) and used the capital

raised to invest in its pipeline, either through internal R&D or by acquiring and/or collaborating with biopharma companies, as dictated by its so-called string-of-pearls strategy.

At \$885 million, the value of the acquisition relative to other early-stage deals—such as the \$2.4 billion BMS paid for Medarex last year (*Nat. Biotechnol.* 27, 781–783, 2009)—seems a bit of a bargain. With big pharma companies eager to get their hands on early- to late-stage drug candidates, why the relatively small price tag? One of the main reasons, according to Krishnan, is that ZymoGenetics, which was founded in 1981, has only one marketed product, and is still very focused on R&D. “ZymoGenetics has no marketing or distribution strength,” says Krishnan. “[BMS] is the best place for Zymo's candidates to get to market.”

Uncertainty about its drug candidates may be reflected in the deal price. PEGylated-IFN- λ , for example, has only completed phase 1 studies, whereas melanoma candidates traditionally have a notoriously high attrition rate. For the pharma company, buying when drug candidates are in phase 2 is significantly cheaper than waiting until after the results of those trials. “BMS will have saved a bit of cash in the long term by making its move now,” says Owide.

Nevertheless, at \$9.75 per share, BMS is paying an 84% premium for ZymoGenetics (at the time of the agreement). BMS spokesperson Jennifer Mauer points out that this premium

is in line with BMS's acquisition of Medarex, Princeton, New Jersey (a 90% premium), as well as other comparable transactions. From the biotech side, however, not all are popping champagne corks. Two days after the deal was announced, at least one disgruntled ZymoGenetics' shareholder filed a lawsuit against the biotech's board of directors alleging they breached their fiduciary duty by selling the company at an unfair price.

The current financial climate caused ZymoGenetics management to do some belt tightening. As well as ceasing its discovery activities in cancer and immunology, in April 2009, the company cut 161 jobs (32% of the workforce) and eight months later, it made another 52 (15%) layoffs to conserve cash. Because of these cash conservation efforts, and a January 2010 public stock offering, Haberman believes “ZymoGenetics is in better financial condition than it was in December 2009.” At the same time, given that most drugs fail in phase 2, the performance of PEGylated-IFN- λ in phase 2 would have become clear only well into 2011 and given the large amounts of cash needed to conduct phase 3 trials, “ZymoGenetics may have felt that the BMS deal was timely indeed,” he says. ZymoGenetics declined to comment to *Nature Biotechnology* as it is currently in a ‘quiet period’ imposed by the US Securities and Exchange Commission.

Emma Dorey Brighton, UK

Shire's replacement enzymes validate gene activation

Genzyme's manufacturing strife and the urgent search for alternatives to meet patients' needs have propelled Shire and its gene-activation technology to the fore. When Genzyme's Allston Landing facility was shut down in 2009 after the discovery of viral contamination, the US Food and Drug Administration (FDA) requested Chineham, UK-based Shire's help in maintaining enzyme supplies for Fabry and Gaucher patients, prompting the company to accelerate its manufacturing timeline for VPRIV (velaglu- cerase alfa, glucocerebrosidase) by 18 months. The reward was an accelerated approval. VPRIV received the FDA's nod in February 2010, and that of the European Medicines Agency in August 2010. Shire's gene-activation technology for generating human proteins has thus emerged as a powerful rival to recombinant technologies. After 20 years in which its potential was obscured by corporate posturing and patent wars, the technology now forms the basis of two of Shire's enzyme replacement therapies: Replagal (agalsidase alfa) for treating Fabry disease, approved in Europe

and currently under fast track designation in the United States, and VPRIV for Gaucher disease.

Gene-activation technology involves introducing a DNA promoter upstream of an endogenous gene in a human cell line. This must be done at a precise location, chosen through knowing the sequence of the gene to be activated. There is an appealing simplicity to the idea of activating an existing gene, rather than the recombinant approach of cloning the gene and introducing it into a mammalian or bacterial cell line. Shire and its antecedents were not alone in pursuing this technology, raising the question of why gene activation has taken so long to translate through to the market.

The answer lies in a fiercely defended intellectual property estate in biotech—Thousand Oaks, California-based Amgen's erythropoietin (EPO) patents—plus a couple of minor skirmishes with Genzyme of Cambridge, Massachusetts, over enzymes for treating Gaucher and Fabry diseases, and lastly, a US Securities and Exchange Commission (SEC) investigation.

Amgen used a monolith of patent rights on its method for manufacturing EPO, to stop Shire

Genetic Therapies' forerunner, Transkaryotic Therapies (TKT; formerly of Cambridge, Massachusetts), from producing EPO. TKT threw down the gauntlet to Amgen in the prospectus for its initial public offering in October 1996, stating that the first approval for its gene-activation technology would be Dypnepo (epoetin delta), a version of EPO, which it was then developing in collaboration with Hoechst Marion Roussel (later Aventis Pharma) of Kansas City, Missouri.

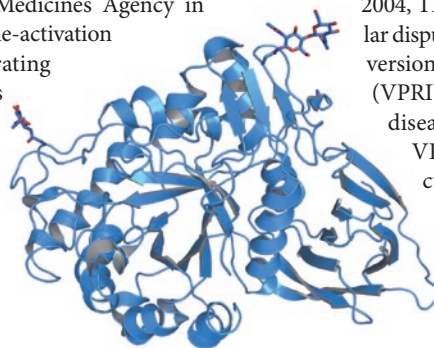
This prompted Amgen to file patent suits against TKT. Although that dispute wound its way around the US and UK courts from 1997 to 2004, TKT found itself in a similar dispute over its gene-activated version of glucocerebrosidase (VPRIV) for treating Gaucher disease. Genzyme claimed VPRIV infringed its cell culture process for manufacturing its marketed treatment Cerezyme (imiglucerase) for the same disease.

At the same time, TKT was in a head-to-head race with Genzyme for the coveted prize of US Orphan Drug status—and seven years'

marketing exclusivity—in Fabry disease. In the event, in January 2004, Genzyme's Fabrazyme (agalsidase beta) received the nod from the FDA, whereas TKT was told that although its drug was approved in 27 other countries, the file for Replagal, its Fabry disease treatment, did not demonstrate efficacy. Just before the FDA reached this conclusion, Richard Selden CEO, who founded TKT in 1988, became the subject of an SEC investigation on suspicions that he withheld information about the FDA's negative views on Replagal from shareholders, while selling shares of his own.

In 2003, Selden resigned from the company and was subsequently found guilty by the SEC in July 2008 and fined \$1.2 million. As the patent disputes trundled on, the new CEO Michael Astrue worked to restore TKT's credibility. Finally, in October 2004, in the third hearing of the case, the UK House of Lords ruled in favor of TKT in the dispute with Amgen.

Little matter that a few days earlier a US federal judge had upheld four of Amgen's EPO claims, ruling that the company had been infringed by TKT, because the UK ruling



Shire generated VIPRIV velaglu- cerase alfa, structure shown here, not by recombination but by targeted activation of an endogenous gene in a human cell line.

IN brief

Pharmacogenomics row



Amgen's Vectibix scrutinized.

A new US government-sponsored report on three pharmacogenetic tests for targeted cancer treatments has confirmed the usefulness

of *KRAS* testing but raised doubts about two other widely adopted tests. The Agency for HealthCare Research and Quality (AHCQR) commissioned the report at the request of the Centers for Medicare and Medicaid Services (CMS) to help set clinical guidelines and reimbursement policies. The report, produced by researchers at the Tufts Evidence-Based Practice Center, Boston, examined published studies linking *KRAS* mutations and the ability to predict responses to two colorectal cancer antibody therapies: Erbitux (cetuximab) from Bristol-Myers Squibb and ImClone, and Vectibix (panitumumab) from Amgen. The team also analyzed studies of genetic variants in *CYP2D6* (cytochrome P450) as response predictors to AstraZeneca's breast cancer drug Nolvadex (tamoxifen); and studies of *BCR-ABL1* mutations as response predictors to Novartis's Gleevec (imatinib) and Bristol-Myers Squibb's Tasigna (nilotinib) in leukemia treatments. The report's conclusion affirmed the value of *KRAS* testing in colorectal cancer therapies but found no evidence for consistent associations between *BCR-ABL1* status and response to tyrosine kinase inhibitor treatment or *CYP2D6* polymorphism status and response to Nolvadex. Mark Ratain, director of the Center for Personalized Therapeutics at the University of Chicago and one of the report's official reviewers, expressed outrage at the *CYP2D6* finding. "This [analysis] is a tremendous disservice to taxpayers and patients," he says. "Medco [pharmacy services firm] is testing all their patients and I would never order this drug without a test." Nolvadex is a pro-drug, and "strong laboratory evidence" exists that individuals with certain *CYP2D6* genotypes cannot metabolize it properly, Ratain notes. According to Ratain, the original version of the report did not reference the largest study by far on *CYP2D6* testing (*JAMA* **302**, 1429–1436, 2009) and "lumps bad studies with good ones." The report's conclusions could encourage more widespread use of *KRAS* tests. But Grace Wang from the Centers for Translational and Policy Research on Personalized Medicine at the University of California, San Francisco, says it might make it harder to get reimbursement for *CYP2D6* and *BCR-ABL1* testing. Overall, she notes, "The report really highlights what we don't know" about the benefit and harms of testing. *Malorye Allison*

IN brief

Joint inspections still cool

Regulatory agencies on both sides of the Atlantic, the European Medicines Agency (EMA) and US Food and Drug Administration (FDA), are urging companies to apply to its joint good manufacturing practice (GMP) inspections because, since its launch in August 2009, the program has had a slow uptake. The regulators aim to increase the number of sites inspected and avoid duplication. But the advantages to companies may be elusive. "Biopharma and API [active pharmaceutical ingredients] manufacturers are undergoing audits and inspections, almost weekly, and this auditing burden is likely to increase under new proposals announced by the FDA in June," says Hedley Rees, founder and CEO of Biotech PharmaFlow, a UK-based, supply-chain management company. To qualify for the joint inspection, companies must have submitted marketing authorization applications in parallel to both the EMA and the FDA, or be hosting a single joint routine reinspection. But these requirements are such that the advantages are lost on would-be applicants. "There are also ingrained cultural differences between the FDA and EMA inspections, with wide variations in requirements and interpretation," adds Rees. "Companies perceive that having a joint inspection will simply raise twice as many issues, leading to negotiations with two parties and the production of two separate reports." *Suzanne Elvidge*

Stimulus trickle

Private biotech companies have received only a small fraction of the \$10 billion from the American Recovery and Reinvestment Act (ARRA) of 2009 funds intended for biomedical research. In fiscal year 2010 the National Health Institute awarded \$196 million dollars of stimulus funding to for-profit organizations, representing 4.2% of the total ARRA funding that passed through the National Institutes of Health (NIH)'s Office of Extramural Affairs, this despite the federal investment's goal of promoting innovation and economic growth in the biopharma sector (*Nat. Biotechnol.* **27**, 587, 2009). Ellen Dadisman, managing director of communications at the Washington, DC-based Biotechnology Industry Organization (BIO), reasons that this trend is consistent with the NIH's intention to direct awards toward basic research. As BIO members are usually focused on translational technologies, so it seems reasonable that they would receive a minor part of the stimulus funding, says Dadisman. "It is our hope that there will be more opportunities for translational/company grants," Dadisman adds, stressing that small companies can receive government funding by other channels, such as Small Business Innovation Research and Small Business Technology Transfer schemes. Additionally, the awards were sometimes announced with very short notice, limiting the number of applicants that could be ready in time. The ARRA Challenge Grants, for example, were announced on 4 March 2009, and had an application closing date of 7 April 2009. *Nidhi Subbaraman*

cleared the way for TKT, which had secured marketing approval for Dynepo (amphetamine and dextroamphetamine) in Europe in 2002, to start looking for a partner to commercialize the product. It was this search that led it into the arms of Shire, then a specialty pharmaceutical company, best known for Adderall, a treatment for attention deficit hyperactivity disorder.

In the event, rather than licensing Dynepo, Shire bought the whole of TKT for \$1.6 billion in 2005.

These disputes over whether human proteins produced by gene-activation infringed the rights of those producing versions of the same proteins through recombinant means are now the stuff of biotech lore. One person who experienced the saga from end to end, including giving evidence about the gene-activation technology in court, is Mike Heartlein, vice president of R&D at Shire Genetic Therapies, who joined TKT in 1989.

Looking back from the perspective of having products on the market, Heartlein says, "It is gratifying when you start with a basic technology and see it develop all the way through the laboratory phase, into the clinic and onto the market."

Heartlein says gene activation can potentially turn on any endogenous gene. In the ten genes activated to date, Shire has in each case used the same DNA promoter. The company has applied homologous recombination to develop other promoters, and Heartlein says there is a continuing research program looking to optimize gene expression. To date, however, no other promoter has proven better than the original.

As a result, gene activation has yet to provide clear advantages in terms of manufacturing, Heartlein says. "That was one of the early promises of the technology—it remains that: the promise has not translated, but it may do so as we discover stronger promoters that augment gene expression," he adds. Heartlein is not really aware of any difference between manufacturing proteins in gene-activated and recombinant cell lines because Shire has never done a direct comparison.

Alongside its effort on promoters, Shire has also done extensive work on the human cell line HT10-80 from the American Type Culture Collection, in which it manufactures its proteins. "We have spent five to six years around converting those cells to make them appropriate for processing in bioreactors," Heartlein notes.

Overall, though, as there is little to choose between gene-activated and recombinant cell lines in terms of manufacturing efficiency, Shire is agnostic about which technology it uses; indeed, Elaprase (idursulfase), its

treatment for Hunter syndrome/mucopolysaccharidosis, which was approved by the FDA in 2006, uses recombinant technology in a human cell line.

Proteins generated by means of gene activation may present advantages, however, in terms of safety and efficacy. Apart from having exactly the same amino acid sequence as the natural product, gene-activation products also have identical glycosylation patterns, a property that is expected to have clinical implications. In November 2007, Shire published work showing that Dynepo has less pronounced angiogenic properties than Aranesp (darbepoetin alfa) *in vitro*. The researchers, led by Alan Stitt at the Centre for Vision and Vascular Science, Queen's University, Belfast, concluded this could be associated with the different glycosylation patterns of the two products.

David Buck, analyst at Buckingham Research Group, in New York, points out that Shire's unique way of manufacturing protein drugs for rare diseases "leads to what [appear] to be advantages in terms of shorter infusion times and less immunogenicity."

"In the clinic, particularly for Gaucher, we've not seen antibodies develop against our product, [which are sometimes] seen with Cerezyme," Heartlein says. "That's exactly what you would predict in terms of immunogenicity."

Buckingham Group's Buck believes gene-activation provides Shire with a strategic advantage, at least "to some extent." In addition, he says the boot may now be on the other foot with Shire, as the company may be able "to block generic versions in future, though it's not something they have played up."

Now other approaches to turning genes on and off, or otherwise modulating endogenous genes, are opening up, notes David Sourdive, executive vice president of corporate development at Collectis of Paris, a specialist in genome engineering. For Sourdive, the advantages of such targeted methods over recombinant techniques are robustness and reproducibility. "You know exactly what you are doing. You don't have to deal with hundreds of thousands of copies of genes that may recombine. Using targeting approaches really makes cell lines robust."

Although the science is in place—and the approval of two Shire drugs manufactured through gene-activation represents important progress—the widespread adoption of gene activation approaches in the manufacture of protein drugs is still likely to take a while, not least because of the length of clinical development. "The time to change is always long, especially when cells are being deployed. But the evolution has begun," Sourdive says.

Nuala Moran London

IN brief

Joint inspections still cool

Regulatory agencies on both sides of the Atlantic, the European Medicines Agency (EMA) and US Food and Drug Administration (FDA), are urging companies to apply to its joint good manufacturing practice (GMP) inspections because, since its launch in August 2009, the program has had a slow uptake. The regulators aim to increase the number of sites inspected and avoid duplication. But the advantages to companies may be elusive. "Biopharma and API [active pharmaceutical ingredients] manufacturers are undergoing audits and inspections, almost weekly, and this auditing burden is likely to increase under new proposals announced by the FDA in June," says Hedley Rees, founder and CEO of Biotech PharmaFlow, a UK-based, supply-chain management company. To qualify for the joint inspection, companies must have submitted marketing authorization applications in parallel to both the EMA and the FDA, or be hosting a single joint routine reinspection. But these requirements are such that the advantages are lost on would-be applicants. "There are also ingrained cultural differences between the FDA and EMA inspections, with wide variations in requirements and interpretation," adds Rees. "Companies perceive that having a joint inspection will simply raise twice as many issues, leading to negotiations with two parties and the production of two separate reports." *Suzanne Elvidge*

Stimulus trickle

Private biotech companies have received only a small fraction of the \$10 billion from the American Recovery and Reinvestment Act (ARRA) of 2009 funds intended for biomedical research. In fiscal year 2010 the National Health Institute awarded \$196 million dollars of stimulus funding to for-profit organizations, representing 4.2% of the total ARRA funding that passed through the National Institutes of Health (NIH)'s Office of Extramural Affairs, this despite the federal investment's goal of promoting innovation and economic growth in the biopharma sector (*Nat. Biotechnol.* **27**, 587, 2009). Ellen Dadisman, managing director of communications at the Washington, DC-based Biotechnology Industry Organization (BIO), reasons that this trend is consistent with the NIH's intention to direct awards toward basic research. As BIO members are usually focused on translational technologies, so it seems reasonable that they would receive a minor part of the stimulus funding, says Dadisman. "It is our hope that there will be more opportunities for translational/company grants," Dadisman adds, stressing that small companies can receive government funding by other channels, such as Small Business Innovation Research and Small Business Technology Transfer schemes. Additionally, the awards were sometimes announced with very short notice, limiting the number of applicants that could be ready in time. The ARRA Challenge Grants, for example, were announced on 4 March 2009, and had an application closing date of 7 April 2009. *Nidhi Subbaraman*

cleared the way for TKT, which had secured marketing approval for Dynepo (amphetamine and dextroamphetamine) in Europe in 2002, to start looking for a partner to commercialize the product. It was this search that led it into the arms of Shire, then a specialty pharmaceutical company, best known for Adderall, a treatment for attention deficit hyperactivity disorder.

In the event, rather than licensing Dynepo, Shire bought the whole of TKT for \$1.6 billion in 2005.

These disputes over whether human proteins produced by gene-activation infringed the rights of those producing versions of the same proteins through recombinant means are now the stuff of biotech lore. One person who experienced the saga from end to end, including giving evidence about the gene-activation technology in court, is Mike Heartlein, vice president of R&D at Shire Genetic Therapies, who joined TKT in 1989.

Looking back from the perspective of having products on the market, Heartlein says, "It is gratifying when you start with a basic technology and see it develop all the way through the laboratory phase, into the clinic and onto the market."

Heartlein says gene activation can potentially turn on any endogenous gene. In the ten genes activated to date, Shire has in each case used the same DNA promoter. The company has applied homologous recombination to develop other promoters, and Heartlein says there is a continuing research program looking to optimize gene expression. To date, however, no other promoter has proven better than the original.

As a result, gene activation has yet to provide clear advantages in terms of manufacturing, Heartlein says. "That was one of the early promises of the technology—it remains that: the promise has not translated, but it may do so as we discover stronger promoters that augment gene expression," he adds. Heartlein is not really aware of any difference between manufacturing proteins in gene-activated and recombinant cell lines because Shire has never done a direct comparison.

Alongside its effort on promoters, Shire has also done extensive work on the human cell line HT10-80 from the American Type Culture Collection, in which it manufactures its proteins. "We have spent five to six years around converting those cells to make them appropriate for processing in bioreactors," Heartlein notes.

Overall, though, as there is little to choose between gene-activated and recombinant cell lines in terms of manufacturing efficiency, Shire is agnostic about which technology it uses; indeed, Elaprase (idursulfase), its

treatment for Hunter syndrome/mucopolysaccharidosis, which was approved by the FDA in 2006, uses recombinant technology in a human cell line.

Proteins generated by means of gene activation may present advantages, however, in terms of safety and efficacy. Apart from having exactly the same amino acid sequence as the natural product, gene-activation products also have identical glycosylation patterns, a property that is expected to have clinical implications. In November 2007, Shire published work showing that Dynepo has less pronounced angiogenic properties than Aranesp (darbepoetin alfa) *in vitro*. The researchers, led by Alan Stitt at the Centre for Vision and Vascular Science, Queen's University, Belfast, concluded this could be associated with the different glycosylation patterns of the two products.

David Buck, analyst at Buckingham Research Group, in New York, points out that Shire's unique way of manufacturing protein drugs for rare diseases "leads to what [appear] to be advantages in terms of shorter infusion times and less immunogenicity."

"In the clinic, particularly for Gaucher, we've not seen antibodies develop against our product, [which are sometimes] seen with Cerezyme," Heartlein says. "That's exactly what you would predict in terms of immunogenicity."

Buckingham Group's Buck believes gene-activation provides Shire with a strategic advantage, at least "to some extent." In addition, he says the boot may now be on the other foot with Shire, as the company may be able "to block generic versions in future, though it's not something they have played up."

Now other approaches to turning genes on and off, or otherwise modulating endogenous genes, are opening up, notes David Sourdive, executive vice president of corporate development at Collectis of Paris, a specialist in genome engineering. For Sourdive, the advantages of such targeted methods over recombinant techniques are robustness and reproducibility. "You know exactly what you are doing. You don't have to deal with hundreds of thousands of copies of genes that may recombine. Using targeting approaches really makes cell lines robust."

Although the science is in place—and the approval of two Shire drugs manufactured through gene-activation represents important progress—the widespread adoption of gene activation approaches in the manufacture of protein drugs is still likely to take a while, not least because of the length of clinical development. "The time to change is always long, especially when cells are being deployed. But the evolution has begun," Sourdive says.

Nuala Moran London

Transgenic salmon inches toward finish line

A fast-growing Atlantic salmon developed by AquaBounty Technologies is poised to become the first transgenic animal to enter the food chain. After a ten-year wait, officials at the US Food and Drug Administration (FDA) reviewed the transgenic fish owned by the Waltham, Massachusetts company. This is the first time the agency has evaluated a transgenic animal destined for dinner tables, though it has done so not as a new food product but as a veterinary drug. Approval hinges mainly on safety of the gene construct for the fish, its effectiveness and food safety. However, some members of the US Congress, experts who advise FDA, environmental groups and activists, and plenty of surveyed consumers have expressed either uncertainty or outright opposition. Many are insisting on labeling of the fish if approved.

"This technology holds incredibly great promise for [boosting] the world's food supply," says Bernadette Dunham, who directs the FDA Center for Veterinary Medicine. Despite that favorable take on the technology, the regulatory review of the fish is continuing, with no decision rendered.

FDA officials made scrupulous efforts to explain the regulatory hoops through which the AquAdvantage salmon had to leap. But to consumer and environmental groups, activists and sundry other critics of this transgenic salmon, the FDA review process proved unpersuasive to some, and outright fishy to others. Qualms over the regulatory evaluation have also been voiced by at least ten US Senators and several members of the US House of Representatives. "There are a number of serious concerns with the current approval process and many potential human

health and environmental risks that are associated with producing GE [genetically engineered] fish have not been fully or openly reviewed," they noted in a letter to FDA commissioner Margaret Hamburg. The members of Congress are insistent that the new transgenic species undergo a formal evaluation by the FDA's Center for Food Safety and Applied Nutrition so that its potential health effects on humans be assessed.

The transgenic Atlantic salmon, which AquaBounty constructed in 1989, contains a gene encoding growth hormone from Chinook salmon, a promoter from another fish species called Pout, and a gene terminator. That construct enables the fish to grow twice as fast selectively bred fish, and that accelerated growth is concentrated into its first year of life. AquAdvantage fish reach 200 g within 200 days, compared with 350 days, on average, for their unmodified counterparts. To reduce the chances that transgenic fish can breed, AquaBounty CEO Ronald Stotish says all the transgenic salmon are female and up to 99.8% are triploid, which renders them sterile.

About 1.5 million metric tons of commercially grown Atlantic salmon are harvested each year, with Chile and Norway the world leaders, followed by the US, the UK and Canada, according to fisheries expert Eric Hallerman of Virginia Polytechnic Institute and State University in Blacksburg. This harvest is mainly pen-grown fish, as there is little remaining wild Atlantic salmon to be found, he says. Aquaculture-raised fish are selectively bred, and grow about twice as fast as their dwindling wild relatives reaching full size in about half the time required for

IN brief

CIRM spurs translation

As the first US Food and Drug Administration-approved clinical trial of human embryonic stem (hES) cells gets underway, the California Institute of Regenerative Medicine in San Francisco (CIRM) is pushing forward with a second round of translational grants. The first round of Disease Team Research Awards, expected recipients to have an approvable investigational new drug (IND) application ready to file within four years. The second round, which will be announced this month, requires that programs have filed an IND or be in phase 1 or 2 by the end of the grant period. The lengthy application process, which includes a six-month planning period, is designed to give teams time to formulate ideas, establish collaborations, and prepare proposals and supporting documentation, according to Bettina Steffen, who, as associate director of development activities at CIRM, oversees the disease teams. CIRM also helps teams set milestones, and evaluate their progress. "We see ourselves as advocates. We want them to put their best foot forward," Steffen says. Geron's hES cell-derived clinical trial is based on work done at the University of California, Irvine, by Hans Keirstead, a CIRM grantee. The trial itself, however, is not funded by CIRM, as its first round of translational grants excluded clinical studies, says Steffen. *Laura DeFrancesco*

Irish bait

The Irish government expects to lure venture capital (VC) firms to its shores with a €500 (\$693) million fund to boost investment in local startups. Innovation Fund Ireland will focus on biotech, information technology, medical devices and cleantech. The exchequer and Ireland's National Pension Reserve Fund will contribute €250 (\$346) million to the fund and interested VC firms are expected to match those contributions. "The Fund is being established to act as a catalyst for an increase in the availability of risk capital for startup and scaling companies," says Garret Murray from Enterprise Ireland, who manages the fund. Dirk Carrez, director for industrial biotech at the European Association for Bioindustries in Brussels, sees advantages in such local initiatives. "National initiatives can be better adapted to regional context and specificities," he explains. They can also be set up much faster than European programs, and "improving access to finance is an urgent problem to be solved for European biotech small and medium-sized enterprises," he adds. The initiative comes on the heels of an announcement in July that a €4.7 (\$6.5) billion Global Pharmaceutical Centre of Excellence (GPCE) is proposed for Tralee, Kerry, Ireland. But funding for the project, promoted by Cork-based generics manufacturer Pharmadel, may be undermined by Ireland's 30 September 'Black Thursday', when the Irish deficit hit 32% of gross domestic product. *Christoph Schmitt*



An AquAdvantage salmon (in the background) and a non-transgenic Atlantic salmon sibling of the same age.

IN brief

Sanofi/Genzyme hostile

Its efforts to acquire Genzyme rebuffed in August, Sanofi-aventis has begun a hostile tender offer for the Cambridge, Massachusetts, biotech, for the \$69 per share (\$18.5 billion) it originally offered. Sanofi notified Genzyme of its intentions in a 4 October letter in which CEO Christopher Viehbacher reiterated that “Genzyme would become the global center for excellence for Sanofi-aventis in rare diseases” and would be managed as a stand-alone division, retaining the Genzyme brand. Genzyme’s board recommended that shareholders reject the offer, labeling it “inadequate and opportunistic” and saying it “fails to recognize the company’s plan to increase shareholder value.” In May, Genzyme articulated a five-point plan, which includes rectifying manufacturing problems, which had led to censure by the US Food and Drug Administration (*Nat. Biotechnol.* **28**, 388, 2010), and disposing of non-core assets including its genetic testing services (sold to LabCorp of Durham, North Carolina, in September). Genzyme also said the offer fails to reflect the value of its pipeline, in particular, the development of its leukemia drug Campath (alemtuzumab) as a “potentially transformative” treatment for multiple sclerosis (MS). Genzyme reported follow-up data from its phase 2 study comparing the drug to high-dose interferon beta 1a, showing that, Campath treatment resulted in lower relapse rates and less increase in disability. *Mark Ratner*

Adverse-events fraud trial

A company’s failure to disclose nonstatistically adverse clinical data does not constitute fraud argues BayBio, the San Francisco-based biotech company association, in an amicus brief submitted to the US Supreme Court. In *Matrixx Initiatives, Inc. et al. v. James Siracusano et al.*, which will be heard by the Court this term, Siracusano alleges that senior executives at Scottsdale, Arizona-based Matrixx misled investors about allegations that its cold remedy Zicam had caused a loss of smell in some patients and that the company’s failure to disclose the complaints led to investment losses. Matrixx claimed that because the adverse-event reports were not statistically significant, the company had no duty to disclose. BayBio COO Jeremy Leffler notes, “Laws requiring disclosure of anecdotal evidence can result in erroneous conclusions about a treatment’s safety and effectiveness,” he says. “As the voice for Northern California’s life science companies, we believe that the laws should require disclosure of significant data collected by organizations.” Matrixx had received several complaints about Zicam from 1999 to 2003, with two doctors compiling data on ten affected patients. But Matrixx officials did not publicly mention the allegations and resulting lawsuits. The US District Court of Arizona granted Matrixx’s motion to dismiss the lawsuit in March 2006. In October 2009, however, the US Court of Appeals for the 9th Circuit reversed that decision rejecting, among other things, the statistically insignificant argument. *Michael Francisco*

pen-grown fish. In the face of that ecological collapse, there are fitful efforts to restore Atlantic wild salmon along with anxiety that escaped aquaculture-raised or transgenic salmon could upset those efforts.

The AquaBounty transgenic salmon are to be bred and hatched at an enclosed facility on Prince Edward Island along the east coast of Canada, and then transported for growth to an inland facility in Panama. Together these facilities offer “better security and reduce the chances for escape,” adding to biological safeguards, including triploid and, thus, sterile, female-only fish, that render them unable to interbreed with wild Atlantic salmon, Stotish says.

“We supply the technology, but don’t want to be the producers,” he continues. Nonetheless, the company envisions transgenic salmon being grown not only in Panama but in confined facilities throughout the United States, yielding local jobs and reducing dependence on imports. This role for transgenic salmon would dovetail neatly with broader growth in aquaculture, an industrial approach that, at 90 tons per year, accounts for about one-half of all fish and seafood consumed worldwide, Stotish says, citing figures from the UN Food and Agriculture Organization in Rome to show that aquaculture is steadily expanding.

Larisa Rudenko, who heads the FDA’s Animal Biotechnology Interdisciplinary Group evaluating the company’s product, points out: “We’re not evaluating the future business plans of AquaBounty.” Instead, the agency sought advice from the Veterinary Medicine Advisory Committee (VMAC) on whether the construct is safe for the fish, effective, safe for consumers, and unlikely to escape or cause problems for wild salmon.

“Any failure of a multiple confinement system means that, once AquaBounty salmon escape, the release cannot be undone because these fish are mobile organisms with very low but not zero likelihood of having some fertile escapees,” says biologist Anne Kapuscinski of Dartmouth College in Hanover, New Hampshire. “It is crucial to conduct a full environmental impact statement [EIS] that assesses the potential genetic and ecological impacts that AquaBounty salmon could have on wild fish and other aspects of the environment. This is even more crucial because of the scientific uncertainty surrounding how these transgenic salmon will function in different environments, the importance of Atlantic salmon as a major global commodity and the existing commitment of US society to restore threatened

and endangered salmon populations and conserve aquatic biodiversity.”

Although no one seriously questions whether these transgenic salmon grow at twice the rate of their wild counterparts, as AquaBounty claims, doubts over safety issues abound. The word ‘preliminary’ came up repeatedly during discussions with VMAC, suggesting that, despite having at least a decade to build a case for approving transgenic salmon, FDA officials and AquaBounty scientists have developed only an incomplete picture on several critical issues.

“This assessment of a genetically engineered salmon...will set a precedent for future approvals of GE animals,” says Michael Hansen, senior scientist of Consumers Union in Yonkers, New York. “Unfortunately, the evidence of FDA’s evaluation of the AquaBounty salmon suggests that FDA has set the bar very low.... This analysis does not conform to FDA standards for assessment of a new animal drug.” He expresses specific safety concerns over potentially heightened allergenicity of the GE fish and poorly executed studies of hormone levels in the fish and their possible health effects on consumers.

Some VMAC members voiced criticisms. “Although I have no particular concerns about the [DNA] construct, it seems inconsistent not to look at the whole profile for safety concerns,” says Gregory Jaffe, the consumer representative on VMAC, who is from the Center for Science in the Public Interest in Washington, DC. Committee member Michael Apley, a veterinarian at Kansas State University in Manhattan, agrees: “We are struggling for a definition of what’s safe, and this is an incredibly important precedent.” For instance, he and others note that there is little information about disease resistance among the transgenic fish. They also have concerns over a condition called jaw erosion that develops in some transgenic—but not unmodified—fish.

The fish are “probably safe” but there are “doubts,” says VMAC chair David Senior of Louisiana State University in Baton Rouge. Senior urged the FDA to include full data sets and to consider conducting a comprehensive EIS. He and others also comment that, if FDA were to approve the AquaBounty fish, the agency should make sure that the company establishes strong site management at both the Panama and Prince Edward Island facilities not only to avoid accidents but also to protect against theft and vandalism.

Jeffrey L. Fox Washington, DC

NEWS maker

Anaphore

This protein engineering firm claims its therapies, modeled on the naturally secreted human serum protein tetranectin, could compete with antibodies.

With the global economy still in the throes of the credit crunch, Anaphore, a company developing a new class of protein pharmaceuticals that it calls 'atrimers' sailed through. In May 2009, the La Jolla, California-based firm raised \$38 million in a series A round from London-based Apposite Capital; GlaxoSmithKline's venture arm SR One of Conshohocken, Pennsylvania; Merck Serono Ventures, headquartered in Geneva; and Aravis Venture Associates of Zurich.

The ready flow of funds from biotech investors points at the appetite for new therapeutic formats that potentially overcome some of the limitations of antibodies. The atrimer technology was developed by Borean Pharma, a spin-off from Aarhus University in Denmark. The Danish scientists created these second-generation, adjustable scaffolds that share the exquisite specificity of antibodies without the drawbacks of size, glycosylation and structure complications that makes manufacture in mammalian tissue culture cumbersome and expensive.

Richard Ulevitch, a professor in the department of immunology of the Scripps Research Institute, initially sniffed out the potential in Borean's technology, in 2007, in collusion with Andrew Schwab, founder and managing partner at 5AM Ventures in Menlo Park, California, for whom Ulevitch served as an advisor. The company was seeded initially by 5AM Ventures, and subsequently by Versant Ventures, both of Menlo Park, California, for a total seed round of \$8 million.

In preparation for acquiring all of Borean's assets, Ulevitch and Schwab asked Kathy Bowdish, now the biotech's CEO, to help with the due diligence. Bowdish, then president of Alexion Pharmaceuticals, in Cheshire, Connecticut, was well placed to see how Borean's work might compete in the second-generation scaffold market. She found herself captivated by the technology. When she expressed interest in coming on board as CEO, Borean's assets were moved to La Jolla in 2008 and Anaphore was born. Founders Ulevitch, Schwab and Bowdish were joined by Phyllis Whiteley, then entrepreneur-in-residence at 5AM. More recently, Russell Greig, formerly president of GSK's SR

One corporate venture group, joined the company as executive chairman.

Anaphore's atrimers are modeled on a human protein of trivalent structure—tetranectin—which is naturally secreted in plasma although its precise function remains unclear. Each of tetranectin's three binding domains contains five amino-acid loops that can be tweaked to bind virtually any target of interest, and Anaphore has systematically changed them, to produce a library of more than 10^{11} versions of atrimers.

Atrimers' three sites engage and 'lock on' to their target with increased avidity, and in theory they could bind any target protein. But atrimers are particularly suited to interact with trimeric targets, a class that includes several proteins of therapeutic interest in the human immune system: tumor necrosis factor (TNF), the receptor for nuclear factor κ B ligand (RANKL) and the TNF-related apoptosis-inducing ligand (TRAIL). But Anaphore's first lead compound, ATX3105, targets the heterodimeric interleukin-23 receptor complex (IL-23R). IL-23R activation normally leads to inflammation, but Bowdish says ATX3105, by blocking soluble IL-23 from docking with its receptor and inducing its subsequent dimerization, can dampen this harmful response in mice.

With molecular weights of 60–70 kDa, atrimers are much smaller than antibodies and achieve greater target tissue penetration. Because they are encoded by one gene and are not glycosylated, they can also be mass produced in *Escherichia coli*. Bowdish says the company doesn't expect any adverse immunological reactions because tetranectin should not be recognized by the human immune system, and tests in mice suggest that the animals don't mount a response against injected tetranectin.

Over 50 companies are currently exploring alternative protein scaffolds, including anticallins (derived from the lipocalin family, featuring an eight-stranded β -barrel structure), adnectins (based on an extracellular domain of fibronectin III) and affibodies (adapted from a three-helix bundle domain of protein A from *Staphylococcus aureus*).



Katherine Bowdish, Anaphore CEO and founder.

Bowdish notes that since acquiring the atrimer platform from Borean, the company's team of 12 scientists has substantially re-engineered and fine-tuned the process, creating a number of new in-house libraries. They use directed evolution to refine initial lead atrimers, using an iterative process of mutation and binding assays. Bowdish says that the five loops offer a large footprint for tinkering and refinement. She adds that the atrimers have naturally occurring lysine-rich unstructured regions at each tip that can be used to conjugate payloads, an option which the company is also exploring.

Arne Skerra, professor and chair of the department of biological chemistry at Technische Universität, Munich, and an expert on emerging protein scaffold technologies, including the anticallins, thinks that the trimeric quaternary structure of the atrimers is an interesting approach that will probably be advantageous when applied to trimeric receptors. He also notes that the avidity of the trimer should be a plus. He is cautious, however, about whether they will be as good as antibodies against nontrimeric target proteins, citing symmetry-related restrictions to the shape space that are imposed by trimers. Anaphore needs to publish some data, Skerra says, before the field can judge its technology.

In the meantime, Anaphore is focusing on autoimmune diseases and secondarily on cancer. Bowdish says that the company aims to bring one or two products to clinical trials unaided, then will collaborate with others outside of the area of immunology and oncology. She predicts that investigational new drug applications from the FDA for a clinical trial could be submitted as early as mid-2012 for an immunological therapy, and late 2012 for an oncological one.

Jennifer Rohn London

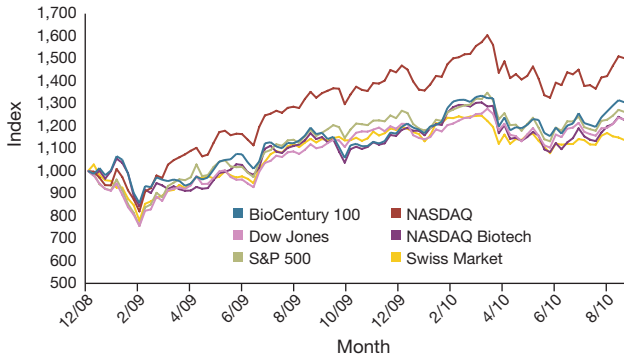
Biotech rallies in Q3

Walter Yang

Biotech stocks stormed back with the rest of the markets last quarter. Although funding for public biotechs was \$9.1 billion, up 55% from the same quarter last year, this figure was largely the result of debt deals completed by Amgen, Gilead and Valeant. For the sector as a

Stock market performance

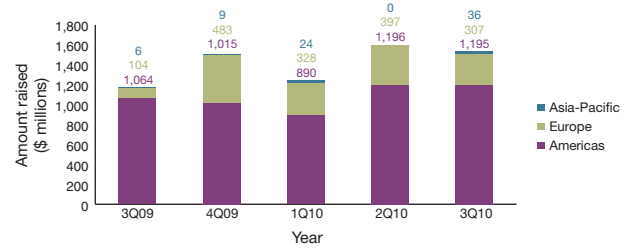
In the past quarter, biotech indices performed (up 12%) similarly to the Dow and S&P 500 (up 10% and 11%, respectively).



whole, equity financings for public biotechs were the lowest since the depths of the economic meltdown in Q408, when the industry raised only \$270.6 million. Private financing remains buoyant (above \$1 billion).

Global biotech venture capital investment

Venture money raised was up 31% to \$1.5 billion from \$1.2 billion 3Q09.

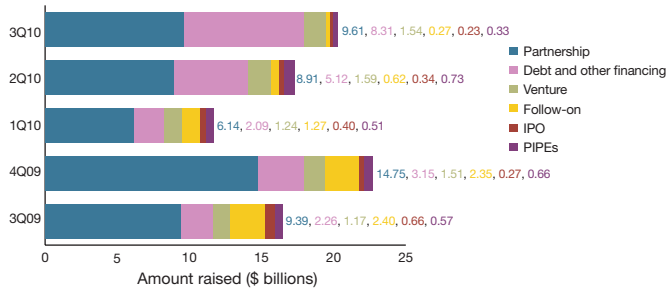


	3Q09	4Q09	1Q10	2Q10	3Q10
Americas	49	58	58	75	52
Europe	14	34	30	27	22
Asia-Pacific	1	1	1	1	3

Table indicates number of venture capital investments and includes rounds where the amount raised was not disclosed. Source: BCIQ: BioCentury Online Intelligence

Global biotech industry financing

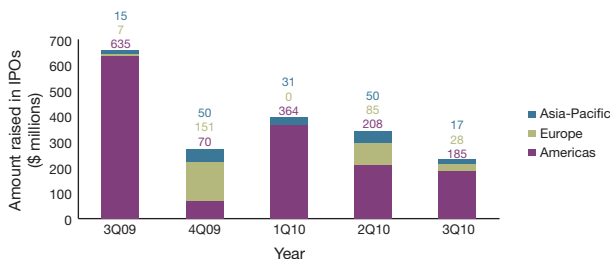
Although money raised by public offerings and PIPEs was down 77% on the same quarter last year, debt financings grew nearly threefold.



PIPEs, private investment in public equities. Partnership figures are for deals involving a US company. Source: BCIQ: BioCentury Online Intelligence, Burrill & Co.

Global biotech initial public offerings

Money raised from IPOs fell 65% to \$230.7 million from 3Q09.



	3Q09	4Q09	1Q10	2Q10	3Q10
Americas	2	2	4	4	3
Europe	1	2	0	5	1
Asia-Pacific	1	2	2	1	1

Table indicates number of IPOs. Source: BCIQ: BioCentury Online Intelligence

Notable Q3 deals

Venture capital	Amount raised (\$ millions)	Round number	Date closed
Company (lead investors)			
Pacific Biosciences ^a (Gen-Probe)	109.0	6	14-Jul
Reata (CPMG, Novo A/S)	78.0	7	12-Jul
immatics (MIG Funds, AT Impf, dievini)	70.5	3	21-Sep
Hopp BioTech			
Relypsa (OrbiMed)	70.0	2	13-Sep
Cerenis ^b (Fonds Stratégique d' Investissement)	51.7	3	26-Jul

Mergers and acquisitions

Target	Acquirer	Value (\$ millions)	Date announced
Talecris	Grifols	3,400	7-Jun
Crucell ^c	Johnson & Johnson	2,200	17-Sep
ZymoGenetics	Bristol-Myers Squibb	885	7-Sep
Movetis	Shire	560	3-Aug

IPOs

Company (lead underwriters)	Amount raised (\$ millions)	Change in stock price since offer	Date completed
Amyris (Morgan Stanley, JP Morgan, Goldman Sachs)	84.8	8%	28-Sep
Seegene	17.5	25%	13-Sep
NuPathe (Leerink, Lazard)	50.0	-28%	6-Aug
Trius Therapeutics (Citigroup)	50.0	-21%	3-Aug
Novagali Pharma (Bryan, Garnier)	28.4	3%	21-Jul

Licensing/collaboration

Researcher	Investor	Value (\$ millions)	Deal description
Arena	Eisai	1,370	Exclusive US rights to market Arena's obesity compound lorcaserin
Aileron	Roche	1,125	Discover, develop and commercialize therapeutics against up to five undisclosed targets using Aileron's stapled peptide technology
Orexigen	Takeda	1,050	Exclusive rights to commercialize obesity candidate Contrave in the US, Canada and Mexico
Seattle Genetics	Roche	912	Adds antigens to a 2002 deal that gave Roche unit Genentech rights to use its antibody-drug conjugate technology

^aIncludes \$50 million from Gen-Probe that was announced in June. ^bFirst close of a series C round. ^cFormal proposal announced in October (after end of Q3). Source: BCIQ: BioCentury Online Intelligence

Walter Yang is Research Director at BioCentury



When patients march in

Impatient with the slow pace of clinical research, families of individuals suffering from untreatable diseases are taking matters into their own hands—with some success. Virginia Hughes reports.

October 4 was the deadline for biotech giant Genzyme of Cambridge, Massachusetts, to respond to serious claims made to the US government in a petition from three of the company's customers. The petitioners all have Fabry disease and are asking the US Department of Health and Human Services to revoke the company's exclusive license on the patent for Fabrazyme (agalsidase beta), the only drug approved to treat the rare condition. More than a year earlier, a virus in Genzyme's manufacturing plant had tainted production of the \$250,000-a-year drug, forcing the company to ration it. Management first estimated that supply would be back on track within one month. This then became six months. And then—after finding particles of steel and rubber in drug vials—the timelines for full production became even longer.

Genzyme now says it expects to stop rationing Fabrazyme in the first half of next year. As for the petition, after reviewing responses from Genzyme, the government will open up the case for 30 days of public comment, and then issue a decision.

In the meantime, the approximately 1,500 individuals who depend on the replacement enzyme have had to make do with only 50% or less of the recommended dose—not enough to ward off the weight loss, burning limbs and kidney distress that come from the disease. Many are skeptical of Genzyme's announcements and feel betrayed. "Genzyme's promises aren't made so much for the patients, but for the investors—they want projections for this quarter or that quarter," notes C. Allen Black, the Pittsburgh patent lawyer representing the three petitioners. But unrealistic and overoptimistic projections from the company, he says, "have really undermined the trust of these patients."

Ironically, Genzyme's first big success, two decades ago, was thanks to fundraising efforts of families affected by Gaucher disease, another rare disorder. Since then, partly because of evaporating venture capital, patient-generated support of the industry has become more common. At least a half dozen companies have been founded or largely funded by patient advocates, and many large nonprofits are funneling grants directly to private companies. "Patient groups are just coming into all stages of product development, helping to meet the [company's] needs and, in many cases, providing resources to do the necessary studies," notes Stephen Groft, director of the Office of Rare Diseases Research at the National Institutes of Health (NIH) in Bethesda, Maryland.

Patient involvement can change the way a company operates—for better or for worse. Advocates push for faster drug development, expedite recruitment for clinical research, and are compelling lobbyists and advocates in government and regulatory settings. But, as the Genzyme shake-up shows, patient and company interests are not always aligned.

Hollywood endings?

Patient advocates carry with them an urgency and passion that makes the toughest of business problems seem almost trivial. Take John and Aileen Crowley. In early 2000, two years after learning that two of their children had a rare and deadly genetic disease, the couple had raised about a million dollars for various nonprofits supporting research into the condition, Pompe disease. But that investment had not led to any treatments, and the children—Megan, then age 3, and Patrick, 2—were getting worse.

Pompe, a glycogen storage disease that disables muscle tissue, affects fewer than 10,000 people worldwide. Children with Pompe usually can't walk or breathe on their own and, at the time, most did not live to adolescence. "Time was short for us," John Crowley recalls.

Along with the ticking clock, Crowley, then a financial consultant at Bristol-Myers Squibb in Princeton, New Jersey, was concerned that his grassroots fundraising efforts, however successful, would not be enough. "It was going to take many tens of millions of dollars to really move the needle, and I didn't know how to raise that for not for profits," he says. "It was time to back one horse."

The 'horse' was William Canfield, a scientist at the University of Oklahoma Health Sciences Center in Oklahoma City, who had just launched a company, Novazyme, to develop a treatment that would replace the enzyme that is defective in Pompe. Crowley became Novazyme's CEO in March 2000. As recounted in the movie *Extraordinary Measures*, while Crowley sought capital investment—initially \$1.2 million from angel investors, followed by \$27 million in venture capital—Canfield's scientific team steadily plugged away at the bench. The following year, Genzyme bought the company for \$137.5 million—one of the largest-ever biotech deals for a drug that had never been tested in humans.

Canfield's treatment was ultimately not developed, but Genzyme's new Pompe program—led by Crowley—did produce a similar compound, which Megan and Patrick received in January 2003. Within 12 weeks, Crowley says he saw "profound effects" in the children, especially Megan, who could suddenly sit up on her own, type, speak and smile. "For both kids, it saved their lives," says Crowley. He is now working on new drugs for Pompe and other rare diseases as chairman and CEO of Amicus Therapeutics, in Cranbury, New Jersey.

For patient advocates, Crowley's story illustrates the primary upside of getting involved with a private company: it's the fastest route to an actual drug. The arrangement can make sense for companies, too, particularly for small startups.

For example, patients played an invaluable role in the early days of Prosenza Therapeutics, based in Leiden, The Netherlands. In 2002, the company was using a new approach—exon skipping—to find treatments for Duchenne muscular dystrophy, a disease that causes rapid muscle degeneration. Duchenne arises from a glitch in one exon of the dystrophin gene. Prosenza's treatments used an antisense oligonucleotide to skip over additional exons, ultimately producing a truncated, yet functional version of the dystrophin protein.



The 2006 documentary *So Much So Fast* chronicles Stephen Heywood's battle with Lou Gehrig's disease, and the company that his brother, Jamie, launched to find a cure. (Source: West City Films, Newton, Mass.)

The first group to show that exon skipping works was led by Gert-Jan van Ommen, head of the department of human genetics at Leiden University Medical Center. Other scientists in the field were skeptical of the technology, van Ommen says, not least because it was sidestepping the root defect of Duchenne. Fortunately, when his grants were rejected elsewhere, van Ommen received a critical infusion of money from the Duchenne Parent Project, a Dutch advocacy group. His results made a big splash in meetings of the close-knit advocacy community. “If people call me the father of exon skipping, then the Duchenne Parent Project was the mother,” he says.

A few years later, ProSensa—whose vice president of drug discovery had been one of Ommen’s postdoctoral fellows—was pushing the science into commercial development. A half dozen Duchenne advocacy groups, including the Duchenne Parent Project, provided funding, in the form of grants, loans and equity. Company representatives declined to say how much money came from advocates, but noted that even a few hundred thousand dollars can be instrumental in a company’s early stages.

“It allowed us to get funded at a stage of the company that you could not get traditional VC [venture capital] money,” says Luc Dochez, chief business officer at ProSensa. Until recently, he points out, few investors saw opportunities in rare disease treatments because they have such small markets. “But because of the personal investment that these patient groups have, they are willing to take financially riskier bets,” he says.

In this case, the bet paid off. In October 2009, ProSensa licensed a treatment for Duchenne, PRO051, to GlaxoSmithKline, of Brentford, UK, which plans to launch a phase 3 clinical trial later this year.

Patient resources

Even after a company is off the ground, patient support can help industry researchers gain a deep understanding of a disease—its symptoms, course, treatment—and what those who have it need most. Dozens of advocacy groups have launched patient registries and tissue banks, which are important to companies during both preclinical and clinical development of drugs (Table 1). For example, if a certain adverse event happens during the course of a clinical trial, it’s crucial to know whether the symptom also occurs in patients who are not taking the experimental drug. “Creating registries, especially in rare diseases, is an area where patients are really key,” says Hans Schikan, CEO of ProSensa.

One well-connected group is the A-T Children’s Project of Coconut Creek, Florida,

which aims to cure ataxia-telangiectasia (A-T), a degenerative genetic disease that affects muscle control, the immune system and ups the risk of blood cancer. Although the disease affects fewer than 500 children in the United States, the organization has raised >\$26 million. In addition to funding research, it has set up a database of families, several blood, cell and tissue banks, and a clinical center at The Johns Hopkins Hospital in Baltimore, where nearly every new case is diagnosed.

Brad and Vicki Margus founded the effort in 1993, shortly after they learned that two of their young boys had A-T. As the research progressed—the gene for A-T (ataxia-telangiectasia-mutated or *ATM*, a member of the PI-3 kinase family) was discovered in 1995—and the organization grew, Brad Margus, who also ran a successful shrimp company, tired of his dual life. “I’d get one phone call from the president of TGI Fridays, and the next from a Nobel laureate,” he says.

In 2000, he sold his food company and launched Perlegen Sciences, a biotech focused on pharmacogenomics and personalized medicine, which Margus ran until 2006. (The company closed its doors in 2009 after mounting operating losses.) He knew that he wouldn’t be able to get investors interested in A-T, because the market was so small. But his new job at least allowed him to stay near the pulse of biomedical research.

At business dinners with executives from big pharma, he had the chance to pick their brains about A-T treatments. In 2006, he started another company, Envoy Therapeutics of Jupiter, Florida, which focuses on common brain diseases, such as Parkinson’s and Alzheimer’s, but also does some work on A-T. “I’m much closer to what I’m passionate about—how to solve a neurological disease,” Margus says.

There are still no effective drugs for A-T, however. The median age of death is 22, and the Margus boys, long confined to wheelchairs, are 21 and 19. Desperate to develop treatments, the A-T Children’s Project supports a few small private companies as well as basic academic researchers. For instance, it is funding a small Arlington, Massachusetts-based pharmaceutical, MindSet Rx, to do animal toxicity studies of one promising compound. The organization is also backing Edison Pharmaceuticals of Mountain View, California, which has a drug that could be tested in A-T children. Clinicians who work at the Hopkins A-T center are helping Edison file the application to the US Food and Drug Administration (FDA), write the trial protocol and recruit patients, Margus says.

“Even more than money, the thing that’s most critical to CEOs of biotech companies is speed,”

Margus says. He adds that the more organized the advocacy group, the more attractive that particular rare disease will be to industry.

Patents and patients

Patient advocates can speed up regulatory and legal hurdles in drug development, too, by using their considerable power in Washington. In early 2000, Patrick and Sharon Terry were making a splash on Capitol Hill—and in certain legal circles—for being the first patient advocates to be involved in the identification of a gene linked to a disease and the control of the intellectual property surrounding it. Like the Marguses and the Crowleys, the Terrys were abruptly pulled into genetic research after learning that two of their young children had a genetic disease, in their case, pseudoxanthoma elasticum (PXE), a connective tissue disease that can cause heart problems and blindness. When they got the diagnoses, just before Christmas in 1995, little was known about the typical course of the disease, and there were few resources for patients and their families.

So, with the help of an umbrella group called Genetic Alliance, the couple launched a non-profit organization called PXE International. Over the next few years, PXE International distributed research newsletters, coordinated scientific meetings and set up a patient registry and biobank.

Patrick was a construction manager who had overseen a lot of the laboratory projects during the Boston biotech boom. After learning of his children’s diagnoses, he visited the laboratory of Klaus Lindpaintner, a researcher at Brigham and Women’s Hospital working on pinpointing the chromosomal region that is disrupted in PXE. Terry asked if he could help around the laboratory, and Lindpaintner agreed. The first day Terry showed up expecting to wash test tubes. “But I was quite thrilled to find that the postdocs viewed me as another collaborator,” he recalls. For several years, he’d go straight from his construction firm to the laboratory, running gels and analyzing data well into the night. Soon Sharon, too, was working in the laboratory.

In 1997, thanks to samples from a family in the PXE International collection, Lindpaintner’s team pinned *PXE* (now called *ABCC6*) to a locus on chromosome 16 (ref. 1). Three years later, thanks in large part to additional samples from the collection, Lindpaintner and three other laboratories isolated the *PXE* gene (and Sharon, in fact, was listed as a co-author)².

The Terrys convinced all of the groups to turn over their patent rights to Sharon who then assigned her rights to PXE International, so that there would be no restrictions put on

Table 1 Selected blood and tissue banks launched by patient groups worldwide

Organization	Resource
Alpha-1 Foundation < http://www.alphaone.ufl.edu/dna_tissue_bank.php >	Set up DNA and tissue bank (in collaboration with the University of Florida)
Association Francaise contre les Myopathies (AFM), France < http://www.afm-france.org/ >	Set up fourteen worldwide DNA and tissue banks
A-T Children's Project < http://www.communityatcp.org/ >	Established patient registry, and cell and tissue banks; identified mutated gene; created animal models; initiated target discovery and compound screening; established CLIA-certified genetic test; developed clinical end points and standardized scale; orchestrated clinical trials.
Cardio-facio-cutaneous (CFC) International < http://www.cfcsyndrome.org/ >	Developed DNA biobank pivotal in identifying components of the MAP-kinase pathway mutant in this disorder
European Research Network for Alternating Hemiplegia (ERNAH), Austria < http://www.enrah.net/ >	Established an SME project, epidemiological register contributing to orphan drug development, registry of cases of alternating hemiplegia in children in Europe, expert database of research project profiles to promote collaboration
Friedreich's Ataxia Research Alliance (FARA) < http://www.curefa.org/ >	Set up patient registry
Hereditary Disease Foundation (Huntington's) < http://www.hdfoundation.org/home.php >	Established blood and tissue bank; led research consortium; gene discovery; initiated clinical trials
Parent Project Muscular Dystrophy USA < http://www.parentprojectmd.org/site/PageServer?pagename=nws_index > Duchenne Parent Project < http://www.duchenne.nl/ >	Set up a coalition to pool knowledge and resources to accelerate development of treatments for Duchenne muscular dystrophy, established a clearinghouse for research grants and location of research resources, developed global patient registry, global clinical trial network
Pacyonychia Congenital Project < http://www.pacyonychia.org/ >	Established patient registry and research consortium; involved in drug development, orphan drug status and sponsored IND application; recruited patients; collected patient samples
Progeria Research Foundation, USA < http://www.progeriaresearch.org/ >	Created patient registry, cell and tissue bank, and database of patient medical records
Pseudoexanthoma Elasticum (PXE) International < http://www.pxe.org/ >	Created blood and tissue bank and mutation database

Adapted from refs. 3,4.

CLIA, clinical laboratory improvement amendments; SME, small and medium-sized enterprises; IND, investigational new drug.

future licensing of the gene test. This was a smart move, especially in light of what was happening to advocates of Canavan disease, a degenerative brain disorder. In 1993, using patient tissue donations, researchers at Miami Children's Hospital isolated the gene for Canavan's. But in 1997, the hospital acquired an exclusive patent on the gene test, preventing it from gaining widespread use. (The advocates eventually sued the hospital, which settled to provide royalty-free tests to researchers searching for a cure for Canavan's.)

After years in the advocacy world, the Terrys became experienced lobbyists for research on rare diseases. After the human genome was sequenced, Patrick Terry testified in Congressional hearings in favor of gene patents. Sharon is now the president and CEO of the Genetic Alliance, which frequently lobbies on behalf of genetic testing, health privacy and open access to government-funded research.

Through his advocacy work, Patrick Terry became a recognizable face in the fledgling genomics industry. In 2000, he and four other entrepreneurs launched Genomic Health of Redwood City, California, a successful personalized medicine company that makes predictive genetic tests for cancer. Terry now

runs his own consulting company, Technic Solutions in Chevy Chase, Maryland, and advises more than 80 international businesses in the personalized medicine field.

There is still no cure for PXE. PXE International's registry now includes some 3,800 patients and its biobank holds 10,000 samples. The organization has funded considerable research on the mouse model of PXE and on possible treatments for the macular degeneration that affects so many people with the disease. "I'm quite positive that before my children lose their vision, we'll have a treatment," Patrick Terry says.

Conflict management

The relationships between advocates and industry are complicated and often fraught with conflicts of interest—demonstrated most recently by the patient petition against Genzyme over the rights to Fabrazyme. The development and early clinical trials of the drug were done by researchers at the Mount Sinai School of Medicine in New York, and funded by grants from the NIH. Mount Sinai later licensed its patent exclusively to Genzyme to carry into commercial development.

Because the initial work was funded by federal dollars, the petitioners claim that

the NIH should open up the patent to other companies that might be able to produce the drug. Any such company would have to meet regulatory approval and give Genzyme a 5% royalty on Fabrazyme sales.

Even if the NIH grants the request, however, it's not clear whether it would mean that patients get the drug any faster. Another company would have to obtain FDA approval of its manufacturing facilities (though lawyer Black says that, given the circumstances, the agency might expedite the process).

"Producing enzyme replacement therapies requires very specialized biological manufacturing procedures that are not easily replicated, and the process to get a new manufacturing facility approved takes years to complete," notes Jamie Manganello, director of global patient advocacy at Genzyme. Since 2006, Genzyme has been building a second plant, which is slated for regulatory approval in 2011. The facility "places Genzyme in the best position to ensure a sustainable supply of Fabrazyme for the future," Manganello adds.

Manganello declined to comment more specifically on the petition, but notes that the company always strives to be transparent with patients, particularly in the past year. "In this period of crisis, we've worked very

closely with patient organizations to give us feedback,” she says. “In many ways, I view many of the patient group leaders as advisors to Genzyme.”

Despite anybody’s best efforts, it’s immensely difficult to find treatments for diseases—even when patient advocates have executive control. When Jamie Heywood started the ALS Therapy Development Institute (ALS TDI) in 1999, he had only one goal: to keep his younger brother, Stephen, from dying of amyotrophic lateral sclerosis (ALS) or Lou Gehrig’s disease. He decided that, because there were no strong candidate drugs for ALS, using a typical for-profit structure would not work. “Making money and helping the person you care about are not easy things to do at the same time,” he says.

Instead, he came up with a unique business model, which he dubbed a ‘nonprofit company’: an organization that was run like a company—with a clear vision from the top down—but funded by donors whose goal was not about generating revenue. He only hired scientists from industry because he felt that academic scientists were too wrapped up in developing theories and accumulating publications. ALS TDI would take on risky, noncommercial projects that companies don’t typically invest in, such as validating mouse models and refining high-quality *in vitro* assays.

The Heywoods were fantastic fundraisers, rounding up more than \$10 million in contributions in the first five years. But they spent the money just as quickly. At the beginning, ALS TDI put a lot of effort into gene therapy; later, Stephen received the first stem cell transplant for ALS in the United States (it didn’t improve his symptoms). The Institute’s scientists claimed to screen more potential therapeutics for ALS than all other research laboratories in the world combined.

Although the researchers made a lot of progress in standardizing methods for ALS models, ultimately, none of the experimental treatments panned out. Stephen died in 2006.

Share alike

The year before he died, Stephen, Jamie and their other brother, Ben, began another venture to meet a different unmet need: helping people with day-to-day management of a disease. What was the best way, for instance, to treat anxiety in someone with ALS? How do you deal with excess saliva in a paralyzed patient? “These are questions that the clinical trial architecture wasn’t answering in any way,” says Heywood. “It’s something you can only learn from other patients.”

The brothers launched a website, called PatientsLikeMe (<http://www.patientslikeme.com/>), where patients share their experiences: their symptoms, the effectiveness and side effects of drugs and general quality of life. The information is entered into a web-based interface that classifies disease according to the classical stages used in the clinic. Today, more than 67,000 people have used the site, which offers the service for eight different diseases. All of the information that might be useful to patients is freely accessible. The company sells other services to pharmaceutical and biotech customers, such as design and recruitment of clinical trials, market research for product packaging and analyses of drug safety.

PatientsLikeMe took a for-profit model, Heywood says, because nonprofits don’t have a good track record for collaboration. “In an environment where you don’t have income as your primary objective, and your primary measure of success is moral righteousness, it tends to put people against each other in ways that are less than constructive,” he says.

Marty Tenenbaum, an entrepreneur and melanoma survivor, learned that lesson the hard way. In 2008, nine years after getting his diagnosis, Tenenbaum launched a company called CollabRx. The initial idea was to provide business consulting and systems infrastructure for patient advocates interested in launching ‘virtual biotechs’—organizations that could use the latest medical research to match individuals with rare

diseases to potentially effective, approved drugs or combinations of drugs. The outcome of each of these pairings would be fed back into a central database, so that the data could benefit future patients. Information from one disease could possibly help people with a different disease.

It was a radical new approach, and one that ultimately didn’t work, Tenenbaum says, because foundations—and particularly the researchers on their scientific advisory boards—weren’t interested in signing up. “They saw that any funding diverted to infrastructure and management of projects like this would take away from the money that was available for science,” Tenenbaum says. “So basically everyone said, ‘Great idea, but we’re not ready for it yet.’”

In the past year, Tenenbaum has restructured CollabRx to give personalized medicine advice to cancer patients who have no time to waste. In parallel to this consulting service, later this year the company plans to launch the first phase of Cancer Commons, an online network for sharing case reports, hypotheses and treatment outcomes from research studies.

Despite the hiccups, there seems to be a growing interest in sharing—among patients, industry and government—to accelerate drug development. Genetic Alliance’s Sharon Terry says that one of the primary goals of the alliance is to encourage more companies and researchers to openly share ‘precompetitive’ findings, such as protocols or endpoints of clinical trials, or data about patients who do not respond to drugs.

“We’re interested in increasing the open space, the commons, where we can have precompetitive aggregation of resources, and figuring out how to de-risk that for companies to allow for the acceleration of research,” Terry says.

Virginia Hughes, Brooklyn, New York

1. Struk, B. *et al.*, *Hum. Mol. Genet.* **6**, 1823–1828 (1997)
2. LeSaux, O. *et al.* *Nat. Genet.* **25**, 223–227 (2000).
3. Ayme, S., Kole, A. & Groft, S. *Lancet* **371**, 2048–2051 (2008).
4. Terry, S.F., Terry, P.F., Rauen, K.A., Uitto, J. & Bercovitch, L.G. *Nat. Rev. Genet.* **8**, 157–164 (2007).

Making the leap into entrepreneurship

Randall Schatzman, Mark Litton, John Latham & Jeffrey Smith

Job security in big pharma is not what it once was. For those looking for a new challenge, a group of big company refugees recount their experience starting up a new life science venture.

Recent years have witnessed numerous layoffs in R&D teams at drug developers and especially in big pharma. According to consulting firm Challenger, Gray & Christmas in Chicago, the pharma industry shed 37,265 jobs in the first 8 months of this year compared with 53,004 jobs in all of last year. These cutbacks have left many talented researchers considering next steps, whether seeking employment with another major drug developer, looking to move to a smaller biotech or even creating and running a new venture.

If you're looking to move on—whether by design or circumstance—and the last option had not occurred to you, it's worth careful consideration. One advantage of creating a new company is that you will be your own boss and you'll be pursuing research in an area that you are passionate about. On the other hand, creating a biotech company from the ground takes years of drive and devotion. But with the right attitude and advice along the way, such an endeavor can bring both personal satisfaction and financial return.

Before you even consider this path, though, it is important to understand that the environment associated with building a startup company is very different from that of a major drug developer. There is a great lack of security that you must be able to overcome if you choose the path of creating a startup. As a first-time founder of a startup has commented to us: "It's like nothing you've experienced in your job at big pharma."

Nurtured from a napkin

For us, it all began on the back of a paper napkin in a local microbrewery outside Seattle.

*Randall Schatzman is president and CEO, Mark Litton is chief business officer, John Latham is CSO and Jeffrey Smith is chief medical officer at Alder Biopharmaceuticals, Bothell, Washington, USA.
e-mail: info@alderbio.com*

The four of us had been given notice from Celltech in Bothell, Wash., where we'd worked in the research facility. We were contemplating our next steps, such as who was hiring or where we might find work. It might have been the beverage, but by the end of lunch we were convinced that there was a better way to discover and produce antibodies, and the answer was through yeast. We began thinking about starting our own company.

We were well aware of the complexities of developing biotherapeutics through our work at Celltech, but we believed we could move more quickly as a small organization. This realization might be the same one other people experience when they're considering starting their own companies—today, more than ever, there is a need for smaller, nimble organizations.

This need for speed and focus influenced the decision to create Alder as a small organization with everyone kept close to the strategy. This led to a hard-to-define internal buzz that you don't find in large corporations, and it drove our team to work harder and faster. In small companies everyone has, in the back of their mind, the knowledge that there is a cliff close by; it's invigorating to know you may head over the edge at any point. This isn't an ideal situation for everyone, but if it works for you, you'll find it to be exciting and even energizing. It can certainly drive focus and motivation. It did for us, and it was especially refreshing after working in a bigger organization for so long.

Focus on proving value

After deciding to forgo the pharma life and start a new company, the next question is always where the focus should lie. Obviously, there is no 'one size fits all' answer for this, but the niche must fit the passion and experiences of the founder(s). In addition, it is important to find an area of research that

can bring value quickly and yet still fulfill an important previously unmet need. Investors are typically interested in obtaining a return within a few years, so demonstrating that you are on track early is necessary. An unmet need is an opportunity, because as a startup you can prove yourself by actively pursuing goals that can bring solutions to larger companies' pipelines.

For Alder, the idea for our technology came from a challenge faced by antibody developers during the transition from research to the clinic. A big issue for most antibody programs occurs in manufacturing, as antibody protein levels are very low and not suitable for commercial production. This means considerable amounts of time are spent looking for sequences that express well in mammalian cells, which delays the development process. Mammalian cell culture tanks are also limited in terms of availability and size, creating long lead times because tanks need to be booked months in advance at extreme cost. Then manufacturing the antibodies themselves takes months before the final product is available for release.

This situation was especially noticeable when a Seattle company launched a great new therapeutic but was unable to manufacture large enough quantities to meet the patient demand. As a result, patients ended up in lotteries to receive the drug.

Alder's solution was to come up with a faster, more efficient way to make large proteins, specifically full-length monoclonal antibodies. This was when we turned to yeast-based cultures, which are fast enough to generate a commercial antibody strain in a few weeks and inexpensive enough to do so at a fraction of what it costs to produce mammalian cell cultures. We solved the problem using what at the time was a new expression system: the yeast *Pichia pastoris*. This technology, alongside a method of finding very

Box 1 Making our way

To date, Alder has raised a total of \$67 million through venture capital. Alder closed a Series A round in July 2005, with Sevin Rosen Funds of Palo Alto, Calif., leading the \$11 million round, joined by Ventures West Management of Vancouver, Canada, and WRF Capital of Seattle. In July 2006, Alder closed its \$16 million Series B round with H.I.G. Ventures of Miami leading and being joined by existing investors Sevin Rosen Funds, Ventures West and WRF Capital. Alder's Series C financing closed in December 2007, with a \$40 million round led by Delphi Ventures of Menlo Park, Calif. and TPG Biotech of San Francisco.

Early on, we made it a priority to find investors that were aligned with our long-term goals as an organization. Finding the correct venture capitalist (VC) for your company is crucial to your eventual success because it is the only way that you will receive flowing capital for necessary trials. Persistence is key in this step, and you must understand from the beginning that you will need to give presentations to countless VCs before finding a group that will be the correct fit.

If you are interested in building your company for the long haul, make sure to find a VC that has a track record of being a company builder. These are classical investors who choose to finance people and ideas over long periods of time. Other VCs will focus more on short-term turnaround and near-term points of liquidity.

Every VC will tell you how to proceed based on its outlook, but at some point you have to start believing in your company and your plan for going forward. VC firms have different needs and strategies. If you listen to each group and change your business plan every time, you will lose focus. Use their advice to sharpen the message but stay true to your own business plan.

high-affinity antibodies, allowed us to get our proof of concept in patients in months instead of years.

That was key, because a common challenge as a startup is understanding that you no longer have the time to spend years identifying new targets and validating them through genetics and genomics as you might have had in your big pharma job. Those efforts would add a tremendous amount of time and risk to a small company's development, which is ultimately detrimental.

As a startup, it is important to think about choosing targets that already have been validated and therefore require less risk to develop. These should be targets that will be attractive to larger companies further along in their pipeline, which will position your company to be the one resolving the larger company's issues with an internally discovered clinical program. This is how you prove to be something of value to big pharma.

Selective hiring

We were fortunate to start as a group of four people with unique backgrounds and diverse areas of expertise that contributed to the overall organization. Eventually, we added two senior research associates to broaden the skill set of the company. These early hires are key. It is important to find people with whom you have a symbiotic and trustworthy relationship because everybody has to be united behind the common goal. This may be a big

change for you if you've come from a major drug developer where employees often are chasing many goals at the same time. It was our persistence and camaraderie that got us through the first 18 months before we landed our Series A funding—salaries are often tight and hours are long in the early days of any company. That's especially true in biotech, so find team members that can get fully behind your vision.

We also knew that we would need a strong legal team if we were to be successful. When building a company, you'll need the correct legal foundations in place from the very beginning. This can be overlooked by new entrepreneurs who are focused entirely on the science, and it's not something any of us had experience with when working at Celltech.

It is important to find an intellectual property attorney who is very knowledgeable in your specific space, not just in the industry. This will allow the attorney to ensure that patents are issued correctly and that you have obtained freedom to operate—a crucial step on the road to success for biotech.

Most venture capitalists (VCs) will have intellectual property attorneys to bring into the operation, but you should try to have your own as well if it is in your budget. In addition, find a corporate attorney who has experience with small startup companies. The first steps in structuring the corporation and its ownership are the base upon which everything else is built, so poor choices can

result in serious repercussions down the line. A first-rate corporate attorney may also have connections in the venture investor space and may be able to provide valuable introductions along the way.

We found our corporate attorney, Sonya Erikson, through Venture Law Group. She helped us put legal foundations in place to build our company the right way. These were things we never had to think about at our old jobs, but if you're launching a company, you'll need representation.

Money, money, money

Today, it is very difficult to find initial investors purely along venture lines, which is why we suggest pursuing angel investors and available grants. These alternative sources for capital are very helpful to get things going and generate preclinical data.

Our first capital came from Department of Defense grants, because the department was interested in our ability to make antibodies quickly in yeast. They put out a call for proposals on methodologies for making antidotes to biological warfare agents, and we were awarded a grant and raised our first \$100,000 as a company. The advantage of grants is that they are non-dilutive and do not set a valuation, which is helpful when you do begin getting venture funding. You won't need to convince anyone of your new, higher valuation. The grant meant we could finally pay our scientists, but we knew that we needed to attract venture funding.

Finding investors is a long and grueling process for any startup, but a necessary step all the same. We were fortunate to have a local biotech entrepreneur and successful CEO make our introductions to the VCs for us. Being active in the local biotech community in the early days of a company is important because it helps build relationships that generate these introductions. When raising money, broader is better—most of the time you will meet with people who are not interested, so we met with anyone who would listen to our story. At the same time, don't be afraid to be selective with the VCs that you identify. Just as team members need to get behind the big picture of the company, VCs must also share this vision to keep things running smoothly. Getting a venture group to say "yes" can be as much about timing as their interest in the science and may depend on the stage of their current fund and the venture group's goals.

It took Alder awhile to find the right investors (**Box 1**). Because of the fact that VCs never say no, there can be many follow-on meetings. This can easily become a never-ending cycle.

What can be helpful is having an interested VC or CEO who knows about your quest make a call for you and get straight feedback from anyone who's turned you away. This will help focus your efforts.

Another excellent source for funding is corporate venture funds (for example, SR One from GlaxoSmithKline, Novartis Venture Funds, Biogen Idec New Ventures and Takeda Research Investment). Once you have a lead investor, these groups usually participate without a board seat and are good sources for future partnering.

Small bites

Another challenge faced by those moving into biotech from pharma is knowing how much venture capital funding to take up front. In the early days of a company, VCs have the ability to be very flexible about valuation, often to the company's detriment. The value set when the Series A financing is completed can sometimes come in low and impact how the company is valued in the future. As a result, it is important to be very cautious at this stage. Be sure to seek advice from entrepreneurs that have gone through venture capital financings before. We learned many things from people


who were willing to share their past experiences with valuation, and we appreciated this advice, especially given our ongoing transition from working with a major drug developer to being entrepreneurs.

In addition, it is tempting to take a large amount of money up front. After all, things will be tight early on and the larger numbers seem appealing. However, taking an excessive amount of money for Series A will dilute the ownership, and that means it lessens your return as a founder. Focus on taking just enough to fund value creation over the next two years. Choosing to do a smaller financing round can put you at risk, but the advantage comes in building value and reaping the benefits of increased value in later rounds.

One of our initial advisors told us, "The first million dollars will be the hardest dollars you will ever raise." This statement rang true because we were forced to raise money with almost no data and a limited track record. The reason we were able to overcome this hurdle was our drive to continually

demonstrate progress. Even with limited resources, we were able to keep showing results at each subsequent meeting. This principle has continued through all our financings. Once you can demonstrate results and meet milestones, you gain credibility. It is this factor that attracts investors.

Conclusions

If starting a new company after leaving a large organization seems challenging to you, that's because it is. Identifying a unique, high value-generating area to research and building an entire organization takes considerable time and effort. It is a different mentality for the entire team to be united to move things forward and avoid the cliff looming in the distance. But the rewards are also substantial, allowing you and your team to pursue the science you find most interesting and promising. At the same time, you may even end up making some money—more than you would have had in your 'safe' job at a big pharma company. 

To discuss the contents of this article, join the Bioentrepeneur forum on Nature Network:

<http://network.nature.com/groups/bioentrepeneur/forum/topics>

Grant management skills are critical for young scientists

To the Editor:

In a commentary last April, Singhvi & Sachdev¹ communicate several skills fundamental for career development of young scientists. Although this article was based upon views from a limited pool of post-docs and faculty at Rockefeller University, New York, it offers a valuable analysis of the universal training challenges faced in preparing scientists to embark upon a successful career, and I sincerely hope it will inspire similar interactive discussion at all research training institutions. Although such an article clearly need not, and indeed cannot, comprehensively address all the requisite skills, I think Singhvi & Sachdev inadvertently trivialize the importance of skills in grant management, specifically financial management aspects and effective collaboration.

One ‘bottom-line’ of their analysis was that obtaining funding is the hard part and that managing grants will simply come naturally—as if these skills are either instinctual or so mundane that anyone will learn them. What a recipe for disaster! This *laissez-faire* approach is certainly ill-suited, in my experience, to modern funding schemes exemplified in European funding frameworks and also in the changing US funding environment, particularly with regard to consortia projects. Long past are the days of *carte blanche*—funding access simply for scientific experimentation and discovery. Increasing oversight from funding agencies and institutions answerable to taxpayers, foundation trustees and stakeholders has ushered in a new order of scientist accountability that demands savvy business skills that must be learned.

Whether we like it or not, grant management entails precise accounting, complicated calculations of salary, overheads

and other costs, which are different for almost every grant instrument and auditing. This involves skills, rarely if ever included in graduate programs, from which we generally shield post-docs so they stay focused on science, and which are often anathema to scientists—otherwise why didn’t we opt for more lucrative business careers in the first place? However, we ignore at our peril the paramount importance of grant management to funding agencies and also to our institutes as legal signees on grants. This is reflected in the fact that many proposals now incorporate a significant weighting for grant management plans in the approval process.

Navigating the grant acquisition labyrinth and writing a successful proposal is just the beginning. Research scientists assume accountability not only for science and laboratory performance (e.g., reaching milestones and fulfilling project deliverables, staff and facility administration, punctual reporting to granting agencies, and publication of grant-financed results), but also for maintaining transparent bookkeeping ledgers that can pass audit muster sometimes for decades hence. Mismanagement of grant funding, regardless of one’s best intentions, is a fast and sure way to crash a career with scant leeway for ‘do-overs’. Unfortunately, delegating bureaucratic responsibilities to financial departments is currently not an option at most institutes. Institutional support is too-often motivated by self-interest to squeeze overhead for diminishing operating budgets and exculpate legal responsibility from potential scientist mismanagement. Rather than lifting the burden, this often creates a second layer of financial reporting for scientists within our own institutes as well as to funding agencies. Although most bureaucrats are receptive to constructive suggestions for improving grant

management with mutual benefit, forums for bipartisan discussion are far too uncommon and delegation of primary responsibility for grant management to scientists seems to be here to stay. Indeed, the assumption attributed to faculty in this article—that grant management will take care of itself—ultimately reinforces administrator perceptions that it is a trivial task not warranting implementation of substantive researcher-focused support infrastructure.

Singhvi & Sachdev¹ highlight a general appreciation among faculty and young scientists for new science as a collaborative venture, but scant definition or direction is provided on how to develop and manage collaborations. Collaboration is traditionally motivated by a need to fill a technological or expertise gap concerning one’s own research topic, and selecting and maintaining such individual collaborations is achieved only through effective networking and genial personal skills (which are arguably difficult to learn). Increasingly, funding availability requires developing and managing consortia of collaborators to achieve complex project goals far beyond the historic paradigm, and this poses new challenges for young scientists that have been overlooked. Consortia-driven projects involving multinational, regional and/or multidisciplinary scientific partners provide a built-in mechanism for funding agencies to ensure diversity and scope to proposals—they are *de rigueur* in Europe and are ascendant elsewhere, including the United States. Leading and even responsibly participating in consortia projects exponentially complicates science management obligations because all partners share responsibilities and consequences. Successful participation in consortia research requires skills that can be best learned through experience and exposure. These skills include selecting reliable partners that work well together, and then motivating these partners to deliver project-oriented results, obtaining punctual financial and/or scientific reports and finally ensuring solvency of this documentation to meet funding agency guidelines. Young scientists should be



exposed early on in their careers to grant administration and start slowly by learning as junior partners. Legal bonds within such collaborations mean that deficient training in project management can lead to unforeseen burdens of compensating for weak links in consortia or to being blackballed from future project opportunities because of unskilled performance aside from scientific output.

'Money makes the world go around' is, and always has been, the lifeblood of a successful scientific career, and effective grant management is critical to maintaining a healthy long-term circulation. By no means should the specter of grant management sour interest in pursuing a scientific career. The increasingly complex funding environment does, however, demand that career training be pragmatic and equip young scientists with the science-based as well as mundane multitasking skills needed for a successful career, and this includes management of funding and

collaboration consortia. Mastering these skills early is critical to avoid derailing the root motivation for choosing a science career—doing good, great science and solving scientific problems—with distractions from management duties that are increasingly imposed upon scientists. Granting agencies and research institutions should also take note though that improving their support infrastructure with scientist input to simplify grant management is essential for the long-term sustainability of quality science.

COMPETING FINANCIAL INTERESTS

The author declares no competing financial interests.

Brion Duffy

Agroscope Changins-Wädenswil ACW,
Swiss Federal Research Station, Wädenswil,
Switzerland.

e-mail: duffy@acw.admin.ch

1. Singhvi, A. & Sachdev, P. *Nat. Biotechnol.* **28**, 378–379 (2010).

least controlling) the levels of these epitopes during biotherapeutics development may be beneficial to patients. Furthermore, the approaches described here to monitor α -Gal levels may prove useful in industry for the surveillance and control of α -Gal levels during protein manufacture.

To determine whether CHO cells potentially express an *N*-acetylglucosaminidase 3- α -galactosyltransferase-1 (Ggta1, E.C. 2.4.1.87)^{1,2}, we first sought to identify and, if present, subsequently clone this gene from Chinese hamster cDNA. Based on sequence alignment of mouse, rat, dog and cow Ggta1 orthologs, we were able to clone by PCR the full-length coding sequence for Ggta1 from cDNA pools derived from four different Chinese hamster tissues (brain, kidney, ovary and spleen) and confirm their homology to the rodent orthologs by direct DNA sequencing (**Supplementary Fig. 1**). The ovary and spleen consensus sequences are >99% identical to each other but do exhibit six single-nucleotide polymorphisms, resulting in one conserved amino acid difference. The remaining nucleotide differences between the orthologs are silent mutations. Molecular cloning of the Ggta1 gene from Chinese hamster cDNA also allowed us to clone directly from the CHO cell genome a critical section of the Ggta1 gene, including exons 8 and 9 and the 5-kb intervening intron adjoining them; exon 9 was of particular interest as it comprises the largest exon within the putative catalytic domain of the galactosyltransferase.

To determine whether the cloned gene product has the requisite features to be an active *N*-acetylglucosaminidase 3- α -galactosyltransferase, we then completed a detailed analysis of the protein structure predicted by the gene sequence. A protein-BLAST search of the CHO Ggta1 exon 8–9 (CHOexo8–9) gene product showed highest sequence homology (>90%) with mouse, rat and bovine Ggta1. We further investigated the structural attributes of the putative enzyme through comparison to existing Ggta1 structures. Notably, the bovine form of Ggta1 is the most extensively characterized, including the availability of several X-ray co-crystal structures for both the wild-type and active site mutants of this enzyme in complex with the UDP-Gal donor and acceptor substrate⁶.

The structural model of the CHOexo8–9 gene product was constructed using the SWISS-MODEL interface (<http://swissmodel.expasy.org/>) in the 'automated' mode. The bovine α -1,3 galactosyltransferase (α 1,3GalT) was automatically selected

Chinese hamster ovary cells can produce galactose- α -1,3-galactose antigens on proteins

To the Editor:

Chinese hamster ovary (CHO) cells are widely used for the manufacture of biotherapeutics, in part because of their ability to produce proteins with desirable properties, including 'human-like' glycosylation profiles. For biotherapeutics production, control of glycosylation is critical because it has a profound effect on protein function, including half-life and efficacy. Additionally, specific glycan structures may adversely affect their safety profile. For example, the terminal galactose- α -1,3-galactose (α -Gal) antigen can react with circulating anti α -Gal antibodies present in most individuals¹. It is now understood that murine cell lines, such as SP2 or NSO, typical manufacturing cell lines for biotherapeutics, contain the necessary biosynthetic machinery to produce proteins containing α -Gal epitopes^{2–4}. Furthermore, the majority of adverse clinical events associated with an induced IgE-mediated anaphylaxis response in patients treated with the commercial antibody Erbitux (cetuximab) manufactured in a murine myeloma cell line have been attributed to the presence of the α -Gal

moiety⁴. Even so, it is generally accepted that CHO cells lack the biosynthetic machinery to synthesize glycoproteins with α -Gal antigens⁵. Contrary to this assumption, we report here the identification of the CHO ortholog of *N*-acetylglucosaminidase 3- α -galactosyltransferase-1, which is responsible for the synthesis of the α -Gal epitope. We find that the enzyme product of this CHO gene is active and that glycosylated protein products produced in CHO contain the signature α -Gal antigen because of the action of this enzyme. Furthermore, characterizing the commercial therapeutic protein abatacept (Orencia) manufactured in CHO cell lines, we also identified the presence of α -Gal. Finally, we find that the presence of the α -Gal epitope likely arises during clonal selection because different subclonal populations from the same parental cell line differ in their expression of this gene. Although the specific levels of α -Gal required to trigger anaphylaxis reactions are not known and are likely product specific, the fact that humans contain high levels of circulating anti- α -Gal antibodies suggests that minimizing (or at

exposed early on in their careers to grant administration and start slowly by learning as junior partners. Legal bonds within such collaborations mean that deficient training in project management can lead to unforeseen burdens of compensating for weak links in consortia or to being blackballed from future project opportunities because of unskilled performance aside from scientific output.

'Money makes the world go around' is, and always has been, the lifeblood of a successful scientific career, and effective grant management is critical to maintaining a healthy long-term circulation. By no means should the specter of grant management sour interest in pursuing a scientific career. The increasingly complex funding environment does, however, demand that career training be pragmatic and equip young scientists with the science-based as well as mundane multitasking skills needed for a successful career, and this includes management of funding and

collaboration consortia. Mastering these skills early is critical to avoid derailing the root motivation for choosing a science career—doing good, great science and solving scientific problems—with distractions from management duties that are increasingly imposed upon scientists. Granting agencies and research institutions should also take note though that improving their support infrastructure with scientist input to simplify grant management is essential for the long-term sustainability of quality science.

COMPETING FINANCIAL INTERESTS

The author declares no competing financial interests.

Brion Duffy

*Agroscope Changins-Wädenswil ACW,
Swiss Federal Research Station, Wädenswil,
Switzerland.*

e-mail: duffy@acw.admin.ch

1. Singhvi, A. & Sachdev, P. *Nat. Biotechnol.* **28**, 378–379 (2010).

least controlling) the levels of these epitopes during biotherapeutics development may be beneficial to patients. Furthermore, the approaches described here to monitor α -Gal levels may prove useful in industry for the surveillance and control of α -Gal levels during protein manufacture.

To determine whether CHO cells potentially express an *N*-acetylglucosaminidase 3- α -galactosyltransferase-1 (Ggta1, E.C. 2.4.1.87)^{1,2}, we first sought to identify and, if present, subsequently clone this gene from Chinese hamster cDNA. Based on sequence alignment of mouse, rat, dog and cow Ggta1 orthologs, we were able to clone by PCR the full-length coding sequence for Ggta1 from cDNA pools derived from four different Chinese hamster tissues (brain, kidney, ovary and spleen) and confirm their homology to the rodent orthologs by direct DNA sequencing (**Supplementary Fig. 1**). The ovary and spleen consensus sequences are >99% identical to each other but do exhibit six single-nucleotide polymorphisms, resulting in one conserved amino acid difference. The remaining nucleotide differences between the orthologs are silent mutations. Molecular cloning of the Ggta1 gene from Chinese hamster cDNA also allowed us to clone directly from the CHO cell genome a critical section of the Ggta1 gene, including exons 8 and 9 and the 5-kb intervening intron adjoining them; exon 9 was of particular interest as it comprises the largest exon within the putative catalytic domain of the galactosyltransferase.

To determine whether the cloned gene product has the requisite features to be an active *N*-acetylglucosaminidase 3- α -galactosyltransferase, we then completed a detailed analysis of the protein structure predicted by the gene sequence. A protein-BLAST search of the CHO Ggta1 exon 8–9 (CHOexo8–9) gene product showed highest sequence homology (>90%) with mouse, rat and bovine Ggta1. We further investigated the structural attributes of the putative enzyme through comparison to existing Ggta1 structures. Notably, the bovine form of Ggta1 is the most extensively characterized, including the availability of several X-ray co-crystal structures for both the wild-type and active site mutants of this enzyme in complex with the UDP-Gal donor and acceptor substrate⁶.

The structural model of the CHOexo8–9 gene product was constructed using the SWISS-MODEL interface (<http://swissmodel.expasy.org/>) in the 'automated' mode. The bovine α -1,3 galactosyltransferase (α 1,3GalT) was automatically selected

Chinese hamster ovary cells can produce galactose- α -1,3-galactose antigens on proteins

To the Editor:

Chinese hamster ovary (CHO) cells are widely used for the manufacture of biotherapeutics, in part because of their ability to produce proteins with desirable properties, including 'human-like' glycosylation profiles. For biotherapeutics production, control of glycosylation is critical because it has a profound effect on protein function, including half-life and efficacy. Additionally, specific glycan structures may adversely affect their safety profile. For example, the terminal galactose- α -1,3-galactose (α -Gal) antigen can react with circulating anti α -Gal antibodies present in most individuals¹. It is now understood that murine cell lines, such as SP2 or NSO, typical manufacturing cell lines for biotherapeutics, contain the necessary biosynthetic machinery to produce proteins containing α -Gal epitopes^{2–4}. Furthermore, the majority of adverse clinical events associated with an induced IgE-mediated anaphylaxis response in patients treated with the commercial antibody Erbitux (cetuximab) manufactured in a murine myeloma cell line have been attributed to the presence of the α -Gal

moiety⁴. Even so, it is generally accepted that CHO cells lack the biosynthetic machinery to synthesize glycoproteins with α -Gal antigens⁵. Contrary to this assumption, we report here the identification of the CHO ortholog of *N*-acetylglucosaminidase 3- α -galactosyltransferase-1, which is responsible for the synthesis of the α -Gal epitope. We find that the enzyme product of this CHO gene is active and that glycosylated protein products produced in CHO contain the signature α -Gal antigen because of the action of this enzyme. Furthermore, characterizing the commercial therapeutic protein abatacept (Orencia) manufactured in CHO cell lines, we also identified the presence of α -Gal. Finally, we find that the presence of the α -Gal epitope likely arises during clonal selection because different subclonal populations from the same parental cell line differ in their expression of this gene. Although the specific levels of α -Gal required to trigger anaphylaxis reactions are not known and are likely product specific, the fact that humans contain high levels of circulating anti- α -Gal antibodies suggests that minimizing (or at

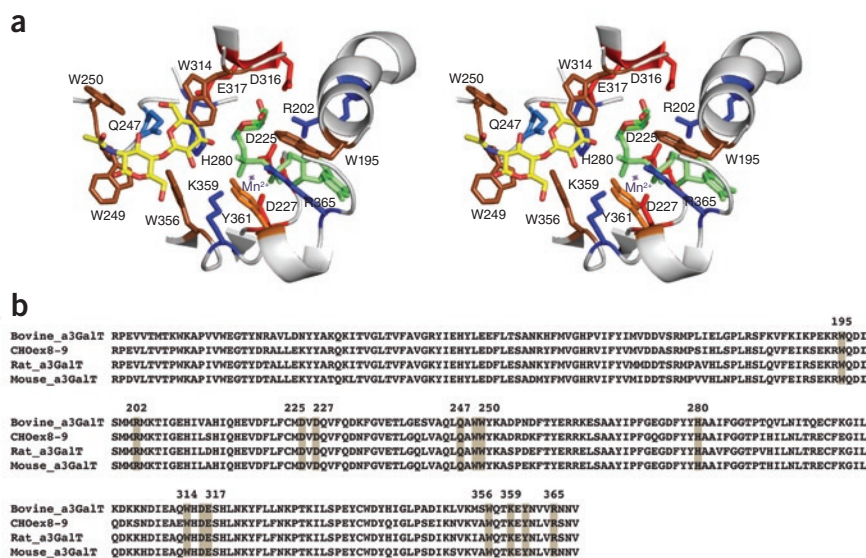


Figure 1 Structural complex of CHO Ggta1 exons 8–9 gene product with UDP-2F-Gal donor and LacNAc acceptor substrate. **(a)** Shown in stereo view is the cartoon representation of the enzyme active site with the side chains of the key residues labeled using the single amino acid code and the numbering based on the bovine α 3GalT crystal structure (PDB: 1G93). The UDP-2F-Gal is shown in stick representation in green and the type II LacNAc acceptor substrate is shown in stick representation in yellow. The location of the Mn^{2+} cation (in purple) is obtained from the bovine α 3GalT crystal structure template used to construct the homology model. **(b)** Sequence alignment of the various α 3GalT with the highly conserved critical active site residues highlighted in gray.

as a template based on its high sequence homology to CHOex8–9. The structural complex of the CHOex8–9 with the donor and acceptor substrate (**Fig. 1**) was constructed with the help of the co-crystal structures of bovine α 1,3GalT with UDP-2F-galactose (UDP-2F-Gal; a donor inhibitor containing a fluorine atom at the 2-position of galactose sugar)⁶ and type II lactosamine acceptor substrate⁷. The critical active site residues in CHOex8–9 are identical to those of bovine α 1,3GalT. Specifically, residues Asp225 and Asp227, along with the oxygen atoms of the diphosphate group of the donor UDP-Gal, are positioned to coordinate a critical divalent Mn^{2+} ion. Furthermore, highly conserved residues Trp249, Trp250, Trp314 and Trp356 are involved in positioning the acceptor substrate for galactosyl transfer. Notably, these residues have been proposed to define the acceptor substrate specificity of the α 1→3 galactosyltransferase family, given that the analogous residues are different in Gb3 synthase, Forssman antigen synthase and human blood group A and B synthases⁸. The presence of these residues in the CHOex8–9 gene product strongly suggests that the cloned gene, if active, is truly an *N*-acetyllactosaminide 3- α -galactosyltransferase.

To confirm the enzymatic function of Ggta1, the catalytic domain of the coding sequence was recombinantly expressed in *Escherichia coli* as a fusion protein possessing an N-terminal maltose binding protein. The expression strategy for the catalytic domain was chosen based on previous work reported for the bovine Ggta1 ortholog⁹. Ggta1 substrate specificity was examined using structurally defined sugar acceptors as substrates including a galactose residue β -1,4 linked to *N*-acetylglucosamine (Gal β 1,4-GlcNAc), *N*-acetyllactosamine (LacNAc) as well as two naturally occurring N-linked glycans: asialo, monogalactosylated biantennary core substituted with fucose (NA2G1F); and asialo, digalactosylated biantennary core substituted with fucose (NA2G2F). The results are summarized in **Supplementary Figure 2** and indicate that Ggta1 is enzymatically active. Ggta1 demonstrates a specific activity of \sim 16 nmol of product $min^{-1} mg^{-1}$ enzyme, using LacNAc as substrate, with an apparent K_M of 200 μ M. Additionally, both N-glycan acceptors NA2G1F and NA2G2F are suitable substrates for Ggta1. Tandem mass spectrometry (MS/MS) fragmentation analysis of the product from the NA2G1F reaction unequivocally demonstrates the specific transfer of galactose from UDP-Gal

onto the terminal galactose of NA2G1F (**Supplementary Fig. 2**); galactosyl transfer onto the terminal GlcNAc is not observed. Furthermore, detailed MS/MS analysis of the enzymatic products strongly suggests the presence of a Gal α 1→3Gal moiety.

To determine the specificity of the Gal–Gal linkage in the product, we then analyzed \sim 150 μ g of the di-galactosylated product (Gal–Gal β 1,4-GlcNA-2AB), generated from LacNAc, by gas chromatography/mass spectrometry (GC/MS; **Supplementary Fig. 3**). In this analysis, terminal galactose and 3-linked galactose are detected as the most prevalent signals observed by GC/MS. To further confirm the stereospecific arrangement of the linkage between the Gal1–3Gal linkage on the Gal1–3Gal β 1,4-GlcNA-2AB product, we subsequently treated the enzymatically generated, digalactosylated product with α -galactosidase and likewise analyzed it by LC-MS. The results clearly demonstrate the susceptibility of the Ggta1-catalyzed product to α -galactosidase (**Supplementary Fig. 3**). These data confirm the specific catalytic functionality of the cloned Chinese hamster Ggta1 gene as an α 1,3GalT necessary and sufficient for the biosynthesis of the α -Gal carbohydrate epitope in a CHO-based cell line.

To determine whether such antigenic species are actually present in pharmaceutical products, we examined in detail a representative biological product produced in CHO cells. Abatacept (cytotoxic T-lymphocyte antigen 4 (CTLA4-IgG)) was chosen because it is a prototypic biologic manufactured in CHO cells, resulting from the fusion of the ligand-binding domain of the CTLA4 and the human immunoglobulin heavy chain constant region.

In abatacept, we observed a glycan species with an abundance of \sim 0.2% of the mass of total N-linked glycans. The unknown species had a neutral mass of 2,070 Da, consistent with the glycan composition HexNAc₄ Hex₆ Fuc₁. Analysis of this structure by MS/MS sequencing yielded fragment ions expected for a typical hybrid glycan structure (**Fig. 2**). However, we observed that the overall fragmentation pattern for this unknown was different from that typically observed for common hybrid structures. Specifically, examination of the ratio between the two B ions with *m/z* (mass-to-charge ratio) of 366 and 528 and the two Y ions with *m/z* of 1,542 and 1,704 suggests that the unknown is more likely to contain an α -Gal structure (**Fig. 2b,c**) than a hybrid structure. Liquid chromatography (LC)-MS/MS analysis before and after treatment with α -galactosidase

showed a clear decrease in the peak containing the 2,070 Da species (retention time ~73.0 min) and a concomitant increase in a glycan with mass of 1,908 Da (retention time ~67.0 min), indicating loss of a single monosaccharide, ostensibly α -Gal. MS/MS analysis on the 1,908 Da product confirmed the expected HexNac₄Hex₅Fuc₁ glycan product (**Supplementary Fig. 4**). Furthermore, quantitative monosaccharide analysis after the treatment of the glycan mixture with α -galactosidase indicated that abatacept contains 370 pmol of terminal α -Gal per milligram of protein (30 mmol/mol of protein; **Supplementary Table 1**). Finally, to confirm and extend these results, we performed LC-MS/MS analysis of glycopeptides derived from a tryptic digest of abatacept, which revealed the presence of an α -Gal containing glycopeptide at one N-linked glycosylation site, namely asparagine 107 (**Supplementary Fig. 5**). The identification of this glycopeptide species demonstrates that the α -galactose glycan is linked to the CTLA4 domain of abatacept and is not derived from a contaminating glycoprotein.

To extend this analysis, the full-length gene encoding Ggta1 was transiently transfected into a previously developed CHO cell line clone that stably expressed CTLA4-IgG. The parental cell line was shown to be negative for expression of the endogenous Ggta1 gene. A highly functional copy of the murine ortholog of the Ggta1 gene was likewise transfected into the same cell line and used as a positive control. Recombinant CTLA4-IgG was purified from the spent media and the levels of α -Gal were determined by monosaccharide analysis as described in the **Supplementary Methods**. The results are summarized in **Supplementary Table 1**. Significantly, CTLA4-IgG expressed from CHO cells transfected with the Ggta1 expression vector contained the α -Gal structure, whereas the control cells mock-transfected with the 'empty' expression vector did not. The amount

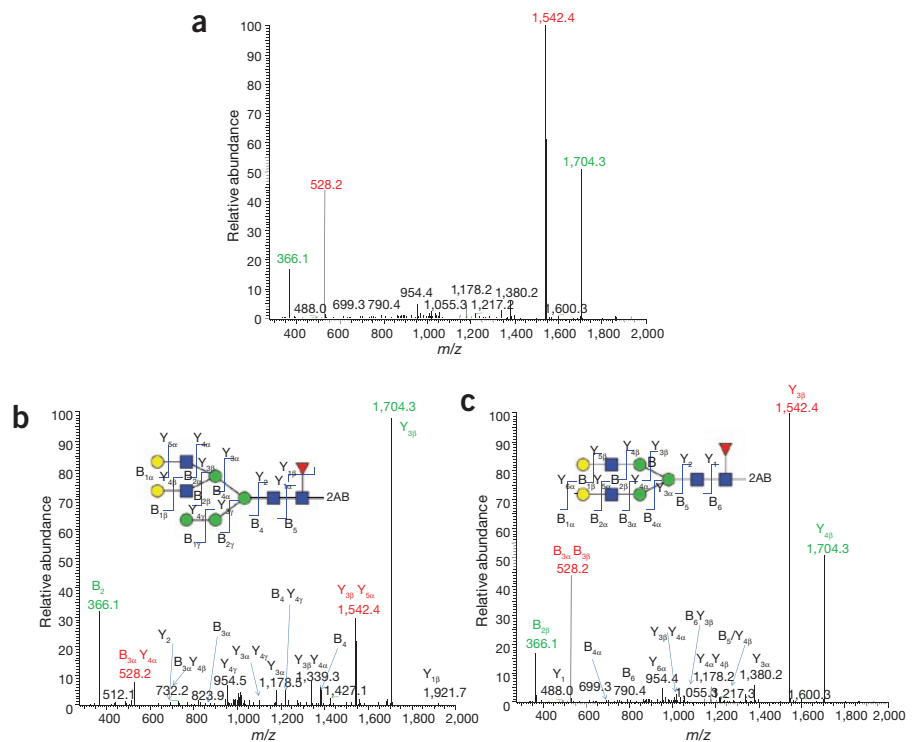


Figure 2 MS/MS fragmentation analysis of N-glycan species observed in abatacept. (a–c) Comparison of the MS/MS fragmentation profile between an N-glycan species with a neutral mass of 2,070 Da observed in Orencea (a) and the fragmentation profile of an isobaric hybrid species (b) and an isobaric α -Gal-containing species (c) from our N-glycan fragmentation database. Nomenclature for glycan structures: blue squares, GlcNAc; red triangles, fucose; green circles, mannose; yellow circles, galactose. The additional galactose is depicted in the lower branch to simplify the fragmentation nomenclature. Fragmentation nomenclature is based on ref. 10.

of α -Gal observed in the recombinantly expressed CTLA4-IgG was comparable to the amount measured for abatacept. Taken together, these results show that commercial biotherapeutics manufactured in CHO can, in fact, contain α -Gal and suggest that the enzyme product of the identified CHO Ggta1 has the appropriate activity to produce α -Gal-containing products.

Given the identification of the gene sequence and to better understand how the α -Gal epitope appears on CHO-based

products, we sought to investigate what cell line backgrounds may have the potential to synthesize the product. To this end, we generated a series of stable clonal cell lines containing the CTLA4-IgG gene from different CHO backgrounds (CHOK1 and CHOdhfr, a line lacking dihydrofolate reductase activity) and used the primers described above to screen for Ggta1 transcript levels (**Table 1**). Interestingly, we observed clones both positive and negative for the Ggta1 transcript, even within the same cell line background, indicating that selection was critical to identify clones with low levels of Ggta1 activity. To correlate the Ggta1 transcript levels with actual output of the α -Gal structure, CTLA4-IgG product was isolated from each clonal population and characterized for α -Gal content (**Table 1**). All clones that expressed Ggta1 at higher levels than background ($C_p < 38$) were also observed to contain α -Gal on the expressed protein product. Furthermore, the level of transcript trended with the level of α -Gal observed on the product. These data illustrate that the expression of the Ggta1

Table 1 Comparison of Ggta1 transcript levels from CHO clones expressing CTLA4-IgG and the corresponding amounts of terminal α -Gal expressed on the abatacept protein product

Clone	Average Cp	Picomol of α -gal/mg protein
Orencea (abatacept)	n.a	369.8
CHOK1 ⁻ clone 0003	31.5 ± 0.4	404.1
CHOK1 ⁻ clone 1145	32.5 ± 0.2	205.4
CHOdhfr ⁻ clone 0301	34.9 ± 0.8	162.7
CHOdhfr clone 1303	38.5 ± 2.1	n.d.
CHOK1 ⁻ clone 1105	35.2 ± 0.4	70.0

Average Cp values for controls (in the absence of reverse transcriptase) were Cp > 38. n.d., not detected.

gene correlates to the expression of the α -Gal product which can arise during the course of clonal selection, and it suggests that any product expressed within these backgrounds has the potential to contain the α -Gal structure.

In conclusion, the data presented here show that, in contrast to what is currently assumed⁵, proteins manufactured in CHO cells may contain α -Gal epitopes. This may be the case especially for proteins with glycosylation sites outside the Fc domains of IgG molecules. The fact that humans contain circulating anti- α -Gal antibodies suggest that controlling the levels of this antigenic epitope during biotherapeutics development could have a positive impact on the clearance and safety profile of the drug. Although the levels of α -Gal in the CHO proteins analyzed in this study are lower in comparison to those typically observed in products derived from murine cell lines, it is important to monitor and control the levels of these species during CHO biotherapeutics development, especially because the specific levels of α -Gal

required to induce anaphylaxis reactions are not well understood and are likely to be product dependent.

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

ACKNOWLEDGMENTS

The Linkage analysis study was supported in part by the National Institutes of Health National Center for Research Resources (NIH/NCRR)-funded grant entitled 'Integrated Technology Resource for Biomedical Glycomics' (grant no. P41 RR018502-01) to the Complex Carbohydrate Research Center. We would like to thank E. Arevalo, S. Wudyka, T. Carbeau and S. Prabhakar for valuable technical assistance and J. Robblee for critical review of 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/>

Carlos J Bosques¹, Brian E Collins¹, James W Meador III¹, Hetal Sarvaiya¹, Jennifer L Murphy¹, Guy DelloRusso¹, Dorota A Bulik¹, I-Hsuan Hsu¹, Nathaniel Washburn¹, Sandra F Sipsey¹, James R Myette¹, Rahul Raman²,

Zachary Shriver², Ram Sasisekharan² & Ganesh Venkataraman¹

¹Momenta Pharmaceuticals, Cambridge, Massachusetts, USA. ²Harvard-MIT Division of Health Sciences and Technology, Koch Institute for Integrative Cancer Research, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. e-mail: ganesh@momentapharma.com

1. Macher, B.A. & Galili, U. *Biochim. Biophys. Acta* **1780**, 75–88 (2008).
2. Larsen, R.D. *et al. Proc. Natl. Acad. Sci. USA* **86**, 8227–8231 (1989).
3. Sheeley, D.M., Merrill, B.M. & Taylor, L.C. *Anal. Biochem.* **247**, 102–110 (1997).
4. Chung, C.H. *et al. N. Engl. J. Med.* **358**, 1109–1117 (2008).
5. Jenkins, N., Parekh, R.B. & James, D.C. *Nat. Biotechnol.* **14**, 975–981 (1996).
6. Jamaluddin, H., Tumbale, P., Withers, S.G., Acharya, K.R. & Brew, K. *J. Mol. Biol.* **369**, 1270–1281 (2007).
7. Boix, E., Zhang, Y., Swaminathan, G.J., Brew, K. & Acharya, K.R. *J. Biol. Chem.* **277**, 28310–28318 (2002).
8. Zhang, Y. *et al. Glycobiology* **14**, 1295–1302 (2004).
9. Shah, P.S., Bizik, F., Dukor, R.K. & Qasba, P.K. *Biochim. Biophys. Acta* **1480**, 222–234 (2000).
10. Domon, B. & Costello, C. *Glycoconj. J.* **5**, 397–409 (1988).

A policy approach to the development of molecular diagnostic tests

Kevin A Schulman & Sean R Tunis

Efficiently generating evidence of clinical utility is a major challenge for ensuring clinical adoption of valuable diagnostics. A new approach to reimbursement in the United States offers a balance between evidence and incentives for molecular diagnostic tests.

Third-party payment through health insurance provides access to new medical technologies for the greatest number of patients by reducing costs to individuals through price negotiation with providers and through direct payment for services. Thus, most technology developers increasingly consider reimbursement decisions by insurers to be as important as regulatory approval for technology development and diffusion. In the United States, this scenario puts scientific development on a collision course with the traditional reimbursement system, as evidentiary standards and coverage policy for this class of technologies remain inconsistently defined. In this article, we propose that a reimbursement pathway—the coverage with evidence development (CED)—first put forward several years ago by the US Centers for Medicare and Medicaid Services (CMS) for experimental products be more fully exploited to facilitate the introduction and refinement of new molecular diagnostics.

Technology assessment and reimbursement

In making decisions about which medical technologies should be considered for payment, both public insurers like Medicare and private insurers determine whether a new medical technology is described in law (for public plans) or contract (for private plans)

Kevin A. Schulman is at the Duke Clinical Research Institute, Duke University School of Medicine and the Health Sector Management Program, The Fuqua School of Business, Duke University, Durham, North Carolina, USA. Sean R. Tunis is at the Center for Medical Technology Policy, Baltimore, Maryland, USA. e-mail: kevin.schulman@duke.edu



In August 2009, CMS issued a final decision that the available evidence does support pharmacogenomic testing for warfarin responsiveness under CED.

as a 'covered benefit' under the insurance program. If the product or service is not a covered benefit, it is excluded from payment under the plan regardless of its merits. For example, federal law precludes Medicare from paying for oral anti-obesity products; Medicare cannot pay for a new therapy without a change in the law, even if the therapy is proven to have a mortality benefit. Detailed descriptions of what is included in an insurance program help to set boundaries on the liabilities of insurers and facilitate the pricing of insurance products.

For those technologies and services that do fit within a covered benefit category, coverage determinations are critical to technology reimbursement and clinical adoption. The Medicare program may not provide payment for products or services that are "not reasonable and necessary for the diagnosis or treatment of illness or injury"¹. Despite sustained effort over many years, no formal definition of this standard has been expounded in regulations². In practice, review by the US Food and Drug Administration (FDA) has been a required first step in the coverage determination process. Even so, differing legal standards and programmatic objectives apply in the approval of a product for marketing and the determination of whether a product is a covered benefit. Not all approved therapies are covered benefits. For example, the devices used for plastic surgery may be approved by the FDA and not covered by a health plan because plastic surgery is excluded from coverage in the insurance contract. Moreover, the FDA may clear a new medical technology for marketing based on similarity to existing approved 'predicate' technologies without reviewing new clinical evidence. This type of FDA review may not provide sufficient evidence to support Medicare payment for those technologies that fit within a potentially coverable benefit category.

Most public and private health plans use some form of technology assessment to provide the scientific basis for making decisions about coverage. In a technology assessment model, these determinations are made based on assessments of available evidence about the proposed product or service. Historically, this process has led to two possible results: first, determination that the evidence supports the

technology, sometimes for specified clinical conditions or patient populations; or second, determination that the evidence does not support the technology.

Applications to diagnostic tests

The first class of widely used clinical applications of molecular biology will likely occur in medical diagnostics, with the potential proliferation of many new diagnostic tools using genomic and proteomic technologies³⁻⁹. These technologies will be introduced into a US healthcare reimbursement system that was not designed (and may not be well suited) to assessing and paying for a new generation of diagnostic testing. The existing system of billing codes for diagnostic tests is poorly suited to the complexity of these new molecular diagnostics, and various proposals are being developed to refine coding mechanisms to address this problem. Equally problematic is the lack of consensus concerning the type of clinical studies that are adequate to demonstrate the clinical utility of these technologies and sufficient to merit their reimbursement. The essence of the issue is whether it is sufficient to demonstrate the sensitivity, specificity and reliability of a new diagnostic test, or whether it is also necessary to empirically demonstrate clinical utility, that is, direct clinical evidence that use of the tests leads to improved health outcomes for patients.

Diagnostic tests, especially molecular diagnostic tests, have been interesting cases for coverage determination. At the extreme, many of the most esoteric diagnostic tests are marketed without FDA review as 'home brew' tests under the regulatory oversight of the Clinical Laboratory Improvement Act (CLIA). Thus, the coverage determination for this category may not even have the benefit of prior FDA review to determine the safety and efficacy of these tests. CLIA standards do not require evidence that diagnostic tests are clinically useful, only that the tests can be performed consistently and accurately.

Although questions about reimbursement are reasonable from a societal perspective, they raise several important corollary issues that will be critical for the scientific community to understand. External validity raises issues of the treatment paradigm, the differences in health practices across markets and the issues of patient population¹⁰. These questions are difficult to address in the development of any particular diagnostic assay, which usually is based on analyses of available data sets that consist of samples from small cohorts of patients and limited information about individual patient characteristics. Clinical utility includes questions of

prognosis versus therapy, treatment impact and outcomes.

Prognosis versus treatment is an issue that arises in the original statute establishing the Medicare program, in which the US Congress stipulated that Medicare would pay only for the diagnosis and treatment of medical conditions, with diagnostic tests defined as those that are provided for patients with existing signs or symptoms of illness, and not in their absence. (Subsequent law has established screening benefits, but usually by listing a single screening test or preventive service in the law rather than adding a broader benefit.) Thus, diagnostic tests that are used primarily for prognostic purposes are not a covered benefit in the Medicare program (and many private payers use similar approaches in their benefit contracts). Treatment impact is an assessment of whether the results of the diagnostic test can be reasonably implemented in clinical practice today. As we define new groups for treatment, we must then assess which treatments work for these new populations we have identified.

Although, in principle, payers might be more excited about tests that can be used to reduce resource consumption by individual patients, there is a concern about the ethics of using the information prematurely. How much evidence would be required to safely withhold a treatment shown to benefit patients in a clinical trial (grade 1A evidence) from patients found to be at low risk for recurrence based on a gene expression profile (currently evaluated as grade C evidence), even if it saves money to do so? Alternatively, the finding that someone is at high risk in a prognostic assay does not define the potential for a patient to respond to more aggressive treatment¹¹. Finally, we are all interested in assuring that individuals have better clinical outcomes as a result of new technologies. However, manufacturers of diagnostic tests have not been held to the high standard of measuring clinical outcomes in a clinical trial. Such a standard has been felt to require too great an investment for the returns available in the traditional diagnostics market, and in some cases, such trials would take many years to complete.

All of these concerns suggest that it will be increasingly difficult to bring new molecular diagnostic tests to patients and increasingly expensive for developers to meet the evidence standard required for reimbursement. Thus, the reimbursement system in its current form may well impede scientific development of a potentially important category of technologies in the absence of thoughtful reforms.

Recent developments in reimbursement policy offer new mechanisms that may promote the adoption and evaluation of promising

technologies such as molecular diagnostics. In 2005, CMS inaugurated a new approach to reimbursement by creating a CED category for technologies that can be reimbursed under conditions of additional investigation into the risks and benefits of the technology¹²⁻¹⁴. The CED policy serves as an announcement that CMS wants to increase access to promising new technologies for Medicare beneficiaries, even when the evidence base for the technology would not satisfy the historical evidentiary requirements that would have reached the level of certainty to be considered "reasonable and necessary" in the diagnosis or treatment of patients¹⁵. CED determinations have been applied to several technologies, including implantable cardioverter-defibrillators for the prevention of sudden cardiac death, biologic agents for colorectal cancer, fluorodeoxyglucose-positron-emission tomography for use in oncology and about a half dozen other technologies. Most recently, CMS issued a CED policy that would provide reimbursement for genetic testing for warfarin sensitivity in an effort to support studies that would determine whether this test would improve outcomes of those requiring anticoagulation therapy.

Applications to clinical biomarkers

As suggested by the recent CMS decision, the CED approach could provide an important pathway for many clinical biomarkers and genomics by encouraging further development of the evidence base on the clinical utility of these new technologies. For biomarkers, the challenge is the time and expense involved in collecting data on clinical utility for each application. Definitive assessments of the clinical utility of strategies that include biologic tests could require large sample sizes and long-term clinical follow-up. For example, the US National Cancer Institute has launched a large randomized trial of a biomarker for early-stage breast cancer and will follow participants for ten years to determine whether women with moderate risk have better outcomes with or without chemotherapy. As the field evolves, this issue may become more complex as multiple additional biomarker assays are discovered, each requiring its own assessment. For private firms developing diagnostic assays, a reimbursement standard that requires high-cost, extended clinical investigations would be equivalent to a statement that they should not pursue these technologies¹⁶.

This is where the CED approach offers an opportunity. The basic concept would involve reimbursement of novel biomarker and genomic technologies on the condition that the patients are included in a data collection effort designed to assess the clinical utility of

the technology. Although the studies themselves would require much thought and care, this approach to reimbursement would result in much more rapid approval, a near guarantee that developers could recoup at least some portion of their investment (for technologies with evidence of benefit) and would provide a foundation for more definitive assessments of clinical utility over time. Ideally, such an approach could be considered across the public and private sectors at the same time.

Issues and opportunities

To implement this framework, we would need to come to consensus on several of the concepts essential to the design of the clinical studies. Specifically, we would need to decide on methods of assessing the clinical validity and utility of these novel technologies. This concept remains the subject of considerable debate, and efforts to achieve these goals have been difficult to achieve in the development of the data collection efforts for the first CED programs. Even so, the investments in health information technology under the American Recovery and Reinvestment Act (ARRA) passed in 2009 might provide key investments in infrastructure to allow this concept to develop. Ideally, to implement this framework we would also develop a common platform for data collection for the multiple potential CED studies occurring simultaneously. This might suggest that the concept of meaningful health information exchange under ARRA would include collection of clinical information for study purposes while being linked to authorization of payment for the specific tests involved. We would want to extend this framework across the public and private sectors as much as possible. The data infrastructure for these assessments could also be the basis for ongoing research by asking

patients for permission to use their biological samples in further research in conjunction with the data collected within the CED framework.

Although payers would be asked to pay for the diagnostic technology, they currently do not pay the costs of the research effort. One advantage of the CED program is that it can be a more efficient way to do the research than creating an independent research study to assess the same outcomes. The costs of clinical care and many of the infrastructure costs are already supported by the claims payment systems (e.g., tracking resource use and outcomes could use the claims data and Medicare demographic files) or by the new medical records technology. The research costs would comprise the costs of the additional clinical data collection required for the study (which would need to be kept to a limited data set), as well as costs of managing the study, aggregating the data and analyzing and publishing the study results. However, the US National Institutes of Health does not have a real-time mechanism to support such efforts, even if the cost reduction were substantial; and Medicare also has no research budget to support such an effort.

In addition, technology companies could be asked to provide funds for the effort as a condition of the CED program, but would likely simply add the costs of the research program to the price of their new products. They might also want influence over the study design commensurate with their level of investment and risk in the project. New public support for comparative effectiveness research could be an additional opportunity to support such a program.

Conclusions

The information required for approval of new diagnostic tests is different from the information that would be required for

optimal decision making in reimbursement environments. The industry cannot support all of these information requirements under the current business model. Our proposal is to consider using the reimbursement framework, and specifically the CED, to provide a bridge allowing early dissemination of new products while assuring that data are collected to prove the effectiveness of the technology. This approach may be particularly important with respect to molecular diagnostics.

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/>

- 42 U.S.C. § 1395y(a)(1)(A).
- Tunis, S.R. *N. Engl. J. Med.* **350**, 2196–2198 (2004).
- Potti, A. *et al. N. Engl. J. Med.* **355**, 570–580 (2006).
- Acharya, C.R. *et al. J. Am. Med. Assoc.* **299**, 1574–1587 (2008).
- Paik, S. *et al. J. Clin. Oncol.* **24**, 3726–3734 (2006).
- Evans, W.E. & McLeod, H.L. *N. Engl. J. Med.* **348**, 538–549 (2003).
- Bild, A.H. *et al. Nature* **439**, 353–357 (2006).
- Rieder, M.J. *et al. N. Engl. J. Med.* **352**, 2285–2293 (2005).
- Liu, R. *et al. N. Engl. J. Med.* **356**, 217–226 (2007).
- Glickman, S.W. *et al. N. Engl. J. Med.* **360**, 816–823 (2009).
- Modan, B. *et al. N. Engl. J. Med.* **345**, 235–240 (2001).
- Anonymous. (USDHHS, CMS, Baltimore, MD, USA) <https://www.cms.hhs.gov/mcd/ncpc_view_document.asp?id=8> (Accessed May 14, 2008).
- Kolata, G. Medicare covering new treatments, but with a catch. *New York Times* (5 November, 2004).
- Tunis, S.R. & Pearson, S.D. *Health Aff.* **25**, 1218–1230 (2006).
- Reed, S.D., Shea, A.M. & Schulman, K.A. *J. Gen. Intern. Med.* **23** Suppl 1, 50–56 (2008).
- Ackerly, D.C., Valverde, A.M., Diener, L.W., Dossary, K.L. & Schulman, K.A. *Health Aff.* **28**, w68–w75 (2009).

What is the value of oncology medicines?

Joshua Cohen & William Looney

Coverage with evidence development (CED), rather than quality-adjusted-life-year (QALY) thresholds, offers the best way forward in balancing evidence-based policy for new oncology products with the needs of developers, payers, physicians and patients.

Faced with balancing increasing demand with limited resources, payers must weigh the costs and benefits of different treatments in deciding which to cover and under what circumstances. Comparative effectiveness research (CER) seeks to bridge the evidence gap between what we know and what we do in medical care by informing this decision-making process. However, potential linkage of CER results to cost-effectiveness thresholds may limit patient access to oncology therapeutics and reduce future investment in pharmaceutical innovation.

In the following article, we provide a counterpoint to claims in favor of the imposition of cost-effectiveness thresholds based on CER and attempt to expose flawed premises underlying the threshold approach. Our aim is to offer a more balanced and nuanced assessment of cancer drugs, including the merits and drawbacks of methods for valuing them and the extent to which such methods constitute a barrier to patient access. Setting thresholds is not a driver of excellent or even efficient care *per se*. Rather than imposing arbitrary limits on drug provision, a more constructive approach to CER-based priority setting is to link reimbursement to the accrual of data that enable further clinical research—as in the case of coverage with CED. Coverage with evidence development (CED) is an evolving method, initiated by Medicare, designed to provide provisional access to novel medical interventions while simultaneously generating the evidence needed to determine whether unconditional coverage is warranted. The CED strategy may serve to reconcile the tension between developing evidence-based

Joshua Cohen is at Tufts University School of Medicine, Center for the Study of Drug Development, Boston, Massachusetts, USA, and William Looney is at Pfizer, New York, New York, USA.



P. Marazzi/Photo Researchers

The benefits of Herceptin (trastuzumab) in early stage breast cancer only became apparent after off label use.

policies for oncology medicines and being sufficiently responsive to product developers, payers, physicians and patients.

Rationing oncology drugs?

Since the mid-1990s, a new wave of molecular, targeted therapies has been introduced to treat various cancers. Although often less toxic than older regimens, they are nearly always more expensive, and on the whole they offer relatively modest gains in survival and quality of life. In a setting of unlimited resources, all effective therapies, even those with modest benefits, could be offered to every patient, who would either pay directly themselves or indirectly through a third-party payer. Because of limited resources, however, publicly funded healthcare programs (e.g., Medicare) and private insurers must weigh the costs and benefits of different treatments in deciding which treatments to cover, and under what circumstances¹.

With the inclusion of \$1.1 billion earmarked for CER in the 2009 stimulus package, as well as

\$500 million to establish the Patient-Centered Outcomes Research Institute under the 2010 health reform bill, the federal government has indicated that it will play a role in this decision-making process. CER seeks to inform decision makers whether and to what extent new technologies offer added benefits when compared with existing therapies². However, because potential linkage of CER results to cost-effectiveness thresholds may have negative effects on patient access to oncology therapeutics as well as future investment in pharmaceutical innovation, we must be aware of potential pitfalls with this approach.

Recently, several oncologists, policymakers and ethicists have criticized what they view as the pervasive prescribing of “marginally beneficial” cancer drugs³. In one particularly provocative commentary, Fojo and Grady⁴ put forward three main recommendations: first, as a matter of policy, payers should not spend more than \$129,090 per QALY, which is the cost-per-QALY of renal dialysis, an amount deemed “sufficient

to buy excellent care"; second, pharmaceutical sponsors should test only those interventions in clinical development that could be marketed at <\$20,000 per treatment cycle if they confer a survival advantage of two months or less; and finally, oncologists should adhere to US Food and Drug Administration (FDA)-labeled indications and avoid as much as possible off-label prescribing of oncology medicines.

In the following sections, we provide a counterpoint to these policy recommendations. Although acknowledging the importance of debate on the critical issue of how much added value new oncology drugs provide, and how CER will contribute to our knowledge base, we attempt to expose flawed premises underlying claims made by Fojo and Grady⁴. We look at the merits and drawbacks of methods for valuing cancer drugs and the degree to which such methods constitute a barrier to patient access. We start by outlining the emergence of health technology appraisals, critically assess the cost-per-QALY approach and then go on to discuss policy implications in the light of our critique.

Health technology assessment in cancer

According to the World Health Organization (Geneva), cancer is poised to overtake heart disease as the world's top killer by 2010 (ref. 5). In the United States alone, 560,000 people died of cancer in 2008, making cancer the second leading cause of death. As such, cancer is a research and development priority, particularly in the United States. Indeed, the United States funds over 60% of global cancer research⁶. In 2005, the average amount *per capita* spent on US cancer research was seven times greater than in the European Union⁷.

Although research spending on oncology therapeutics is relatively high, the success rate (that is, the likelihood that a compound entering clinical development will eventually reach the marketplace) is comparatively low (8%)^{8,9}. Carefully targeted clinical trial design is particularly vital in oncology research, as is finding the appropriate pool of patients. Both challenges contribute to high development costs. Moreover, trials often must enroll the sickest patients, given the potential risks and ethical dilemmas of treating earlier stage, potentially curable, cancer patients with powerful compounds that may have significant or unknown adverse side effects. Relatively high per-unit prices of new oncology therapeutics are partly a function of high development and production costs. Given the high price per treatment cycle of most new anti-neoplastic agents, the vast majority of patients must resort to third-party payment. Public and private payers play the role of intermediary, evaluating newly marketed therapeutics, from both a clinical

and economic vantage point, before approving them for prescribing and reimbursement. This process has been termed the "fourth hurdle," joining the traditional hurdles of safety, efficacy and manufacturing.

The international trend is for public payers to systematically adopt the fourth hurdle. Negative health technology assessment (HTA) appraisals often lead to coverage refusals or conditional reimbursement policies. Conditions may include indication restrictions, prior authorization, quantity limits and step therapy. Finally, in certain instances, HTA drives decisions on what price is acceptable for a new treatment relative to the price of existing therapy.

Fojo and Grady⁴ favor more systematic adoption of the fourth hurdle in the United States, specifically for cancer therapeutics, following the international trend. Moreover, they assign a responsibility to oncologists as "agents of change" to internalize HTA appraisals, whether in the form of payer policies or clinical practice guidelines and limit their prescribing of "marginally effective medicines."

The cost-per-QALY approach

The most prevalent approach to measure cost effectiveness is the cost per QALY. It is the approach explicitly advocated by HTA authorities and implicitly endorsed by Fojo and Grady⁴. In simple terms, a QALY incorporates both the time by which a new treatment extends a person's life and the quality of life the patient experiences during that added time. Here, each year of perfectly healthy life is assigned a value of 1.0, whereas death is fixed at zero. If each year is accompanied by some form of burden that is less than full health, then an intermediate value—a quality-of-life adjustment factor—is assigned. Thus, drug A, which promises an added 10 years, but has a quality adjustment factor of 0.5, due to burdensome side effects, produces 5 QALYs (10 × 0.5). The cost per QALY indicates the incremental cost-benefit ratio of drug A compared with, say, existing drug B. If, for example, drug A costs \$50,000 more than drug B, its cost-per-QALY equals \$10,000.

The rationale for any form of cost-effectiveness analysis is that it offers an explicit approach to quantify the costs and benefits of a technology using a common denominator (e.g., a QALY). The resulting cost-effectiveness ratios can be compared with each other or with a threshold value across a wide range of interventions for different diseases, with the goal of identifying the most efficient way of maximizing health at the population level. In this respect, the QALY represents an average. That is, it does not predict what an individual might experience, but describes what one might expect, on average, from a whole population.

Despite its merits as a decision tool, the cost-per-QALY approach is seen by many experts as a blunt instrument, allowing for "scant sensitivity in comparing the efficacy of two competing but similar drugs"¹⁰. Key flaws in the QALY methodology are:

- An inability to provide a rationale or empirical basis for arbitrarily chosen thresholds, which are selected by payers with no input from patients or oncologists¹¹.
- An inability to accurately account for complexities of measuring individual preferences in capturing the trade-off between quality and length of life¹².
- An assumption that policymakers should maximize QALYs subject to a budget constraint, irrespective of the parties to which QALYs accrue. This contrasts with the views of the general public, which, when asked, appears to want to prioritize resource allocation in such a way that we help those who are worst off first, because they have the greatest unmet needs^{13,14}.
- An ageist tendency, as QALYs place greater value on years gained earlier in life. This imparts an inherent bias toward younger adults, with longer potential for additional years of good health¹⁵.
- And finally, a reliance on data from randomized clinical trials. Payers seek to carry out HTA studies before products are marketed. However, at pre-launch, as well as at the time when a first market license is granted, evidence on the full clinical potential is largely incomplete, and there will be no actual cost data for the new product based on normal health system usage.

Cost-per-QALY estimates pertaining to recently approved oncology drugs have tended to be relatively high¹⁶. As a result, many cancer medicines have not met the HTA thresholds. In fact, between 40% and 55% of oncology medicines approved after 2000 have either been given outright denials or indication restrictions by three major HTA agencies: UK National Institute for Health and Clinical Excellence (NICE; London), the Australian Pharmaceutical Benefit Advisory Committee (Canberra, ACT, Australia) and the Canadian Common Drug Review (Ottawa, ON, Canada)¹⁶. Moreover, evidence from a study of NICE recommendations indicates that of 55 cancer medications assessed between May 2000 and May 2008, nearly half were rejected for routine use in the UK National Health Service (NHS) or restricted to a narrow

set of indications¹⁷. The latter study classified NICE decisions as “positive,” “restricted” or “negative” before and after a change in NICE appraisal methods in August 2006. The change entailed an accelerated process for technology assessment (**Table 1**). Classification relates to the degree to which routine NHS use is recommended for all (positive), some (restricted) or none (negative) of the patient groups included in a drug’s registration. The proportion of ‘negative’ appraisals increased from 4% in period 1 to 27% in period 2, reflecting, in part, the fact that cost effectiveness has become the criterion *de rigueur*.

Thresholds, price controls and restricting off label

Although there are opportunity costs associated with spending on, or choosing, marginally beneficial treatments, we believe that policy recommendations, such as those put forward by Fojo and Grady⁴, are both problematic and unconstructive for several reasons.

The QALY threshold. The majority of newly approved cancer drugs would not pass the arbitrarily chosen renal dialysis QALY threshold, let alone the NICE threshold of ~\$50,000 per QALY. Certainly, some cancer drugs with cost-per-QALY ratios above \$129,000 are worth prescribing, among which may be those that cure a previously incurable condition, or those that produce prolonged remissions with minimal toxicity or those that reduce the amount of time a patient spends in the hospital. With respect to the latter, costly medicines could displace in-patient spending, as appears to be the trend in cancer care, according to a recent comprehensive study on cancer spending¹⁸. Escalating costs of new cancer drugs are receiving a lot of attention. However, as a proportion of overall cancer expenditures, drug spending is only 6% (ref. 18). But more importantly, Fojo and Grady’s policy prescription begs the question, Is limiting spending to less than \$129,090 per QALY necessarily sufficient to purchase excellent care? In many instances it is. However, setting thresholds is not a driver of excellent or even efficient care *per se*.

Apart from the theoretical shortcomings of the cost-per-QALY approach listed above, there are also practical considerations. Studies of pan-European access to, and uptake of, anti-cancer therapies suggest that countries that rely on the cost-per-QALY approach to inform resource allocation decisions tend to implement more restrictions and delays in access¹⁹. Owing to the length of time of appraisals, the HTA process may cause delays between the point at which a product is launched and the decision to reimburse.

Table 1 NICE recommendations on cancer drugs

	Number (percent) of appraisals		
	May 2000 to October 2008	May 2000 to June 2006 (period 1)	July 2006 to October 2008 (period 2)
Positive appraisals	18 (47)	11 (48)	7 (47)
Negative appraisals	5 (13)	1 (4)	4 (27)
Restricted appraisals	15 (39)	11 (48)	4 (27)
All appraisals	38	23	15

Adapted from ref. 17.

Furthermore, in terms of overall effectiveness and outcomes of cancer care, the United States appears to compare favorably to several countries with healthcare systems that ration by way of cost-effectiveness thresholds. For example, the five-year, age-adjusted cancer survival rates for the United States and seven European countries show that the United States has superior five-year survival rates for six different common cancers^{20,21}. Compared with the United Kingdom, for instance, the US five-year survival rate is between 25% and 60% higher for breast, thyroid, prostate, lung, colon and cervical cancers. In certain instances, such as non-small cell lung cancer and breast cancer, where certain drugs have a proven effect on survival, lack of access in the United Kingdom is likely a contributing factor to lower survival rates. At the same time, access restrictions may be compounded by prevailing nihilism and delayed diagnoses, as patients with lung cancer, for example, frequently receive no treatment at all in the United Kingdom^{22,23}. In this context, it is important to note that the incremental cost-effectiveness ratios of Tarceva (erlotinib), Avastin (bevacizumab) and Erbitux (cetuximab) for individuals with advanced non-small cell lung cancer exceed \$129,090 per QALY. Yet, Medicare and some private US payers cover these drugs, and they are correspondingly used more in the United States than in the United Kingdom^{24–26}.

In practice, the cost per QALY threshold delineates a cut-off between technologies payers can afford for all, as opposed to those they will fund for no one. Indeed, it is the premise of the English and Welsh healthcare systems that if they cannot afford a new anti-cancer drug for all who might benefit, the drug will not be paid for. More equity is desirable, particularly a reduction in geographic outcome variation. However, is such a radically egalitarian, one-size-fits-all approach warranted?

Ironically, at a time when some medical experts are recommending the imposition of explicit thresholds in the United States, NICE is moving away from strict reliance on cost effectiveness and thresholds, allowing for other factors to enter the equation, in particular, for cancer therapeutics. In response to pressure

from key stakeholders, NICE is considering the size of the patient population, the degree to which end-of-life treatment is targeting patients with short life expectancy (which implies attaching more value to survival gains when life expectancy is short) and the availability of alternative treatments. To illustrate, in 2008, NICE declined to recommend Sutent (sunitinib) to any patients with advanced kidney cancer²⁷. If followed uniformly by NHS acute and primary care trusts, as is the intended policy, its decision would have resulted in blanket denial of care. However, in 2009, after changes in its stated policy on end-of-life treatments, NICE reversed its decision, stating that sunitinib should be a first-line treatment option for some patients with advanced or metastatic renal cell carcinoma²⁸.

De facto price controls. Some experts also suggest that drug sponsors impose on themselves pre-launch price controls. But determining a drug’s clinical value is something that cannot and should not be decided before a drug’s approval. This is what markets do after approval; what’s more, considerable uncertainty is associated with a drug’s real-world effectiveness. Often, a therapy’s effectiveness becomes apparent only after it has been used in thousands of patients, rather than the tens or hundreds participating in trials. Thus, by introducing a maximum price per treatment cycle for a particular cancer pre-launch, drug developers may consider the cost-benefit ratio (that is, cost of investing in further drug development versus return on that investment) unattractive. This could result in a reduction of innovative drugs entering the clinic for that cancer.

Off-label practices. Severely restricting off-label use stands in stark contrast with clinical practice. Depending on the drug in question, studies have shown anywhere between 4% and 75% off-label use of anticancer agents²⁹. Important benefits are often discovered after a drug is initially approved (that is, the period during which many cancer drugs are prescribed off-label). For example, Herceptin (trastuzumab) was initially approved for end-stage breast cancer; only later was it discovered to have benefits when used at

earlier stages of disease. Also, Taxol (paclitaxel) was initially labeled as a second-line agent for ovarian cancer. Now, it is also indicated for Kaposi's sarcoma, non-small cell lung cancer and breast cancer. A similar story can be told for Gemzar (gemcitabine), Paraplatin (carboplatin) and numerous other cancer agents.

Oncology is a clinical specialty where treatment urgency drives a high prevalence of experimental (that is, off-label) use of drugs beyond their approved indications. This does not imply that evidence should not play a role in establishing the value of, or limits to, their uses. Payers have shown reluctance to reimburse drugs that are not approved by the FDA on the grounds that their use is experimental or investigational³⁰. It would seem they are justified in requiring that healthcare physicians provide evidential support, such as peer-reviewed studies, clinical practice guidelines and officially recognized compendia, for the purpose of off-label use.

The need to address diverse perspectives

In a recently published polemic in favor of explicit rationing, the controversial utilitarian philosopher Peter Singer posits, "You have advanced kidney cancer. It will kill you, probably in the next year or two. [Sunitinib] slows the spread of the cancer and may give you an extra six months, but at a cost of \$54,000. Are a few more months worth that much?... Is there any limit to how much you would want your insurer to pay for a drug that adds six months to someone's life?"³¹ The answer is probably yes. Surely, limited resources require that society address what counts as a benefit, the extent to which cost should factor in deliberations and who should be involved in those decisions. But, the key questions then are, Who decides on such limits and how are limits determined?

There is room for building a more systematic and better coordinated evidence base in the United States for both labeled and off-label use. Accordingly, CER could help close the gap between what we know and what we do in pharmaceutical care. Here, CER should involve genuinely participatory decision-making among all relevant stakeholders, which entails a complexity perhaps overlooked by Fojo, Grady, Singer and other advocates of a hard-nosed, utilitarian approach to CER implementation using cost-effectiveness thresholds. CER must take into account the different (sometimes conflicting) perspectives among the primary stakeholders in cancer care: patients, oncologists, the pharmaceutical industry and payers.

The first perspective is that of a terminally ill cancer patient, who often places a very high value upon extending the years, months or even days of good quality life he or she has remaining, with a high willingness to pay for new cancer

therapies^{32,33}. Yet, the patient may not be able to afford treatment, with or without insurance. A second perspective is that of oncologists, who may find themselves in the unenviable position of having to serve as advocates for their patients' best health interests while being forced to help them make decisions about whether the potential clinical benefits warrant the financial burden. A third viewpoint is that of the drug industry, which needs relatively unhindered access to the market and a price that offers a reasonable return on investment. Capturing that return provides the basis for future investment in innovation. Drug firms facing price controls may be reluctant to invest in future innovations (that is, they may anticipate effects of price controls and curtail research accordingly) as potential payoffs are reduced. A final outlook is that of the payer, whether public or private, who face a formidable challenge in allocating limited resources across many diverse disease areas of acute and chronic medical need. Many seek to evaluate the broader, population-wide, added value (that is, cost-effectiveness) of cancer drugs. Rationally, they have to determine what can be paid for. In this respect, they must be mindful of interventions that may have to be forgone—opportunity costs—in other disease areas if rising numbers of expensive cancer drugs are to be provided to a growing patient population.

A more balanced solution?

We do not hold a binary view on resource allocation—the choice is not to ration in accordance with CER-constructed cost-effectiveness thresholds or allow unfettered access. Rather, we maintain that there are more constructive CER-based approaches to priority setting for oncology therapeutics, such as CED, in which reimbursement is linked to requirements for further clinical research³⁴. Here, CED generates data to be used as the basis for future coverage decisions, which implies the possibility of discontinued reimbursement if findings indicate lack of comparative (cost) effectiveness. CED is often considered in cases where a medicine is initially evaluated as not cost effective, whereas medical (unmet) need is high.

Where appropriate, CED could be combined with risk-sharing arrangements, in which drug makers and payers agree to relate a product's future performance to its price, revenue stream or reimbursement. As such, CED may serve to reconcile the tension between developing evidence-based policies and being responsive to product developers, physicians and patients.

Using the time-honored randomized clinical trial methodology, Medicare has already implemented at least ten CED programs, including the off-label use of colorectal cancer drugs. However, to truly close the gap between what

we know and what we do in the provision of medicine, CED should rely on alternatives to randomized clinical trials, such as observational studies. Furthermore, to adhere to the tenets of patient-centric outcomes research, CED should emphasize personalized (targeted) therapy. In this respect, we believe the recently proposed CED on pharmacogenomic testing for warfarin response is a step in the right direction.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Yabroff, K. & Schrag, D. *J. Natl. Cancer Inst.* **101**, 1161–1163 (2009).
2. Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the Twenty-First Century* (National Academy of Sciences, Washington, DC, 2001).
3. Mulcahy, N. Time to consider cost in evaluating cancer drugs in the United States? <<http://www.medscape.com/viewarticle/705689>> (2009).
4. Fojo, T. & Grady, C. *J. Natl. Cancer Inst.* **101**, 1044–1048 (2009).
5. Anonymous. CDC Faststats (Centers for Disease Control and Prevention, Washington, DC, USA) <<http://www.cdc.gov/nchs/FASTATS/cancer.htm>> (Accessed, November 1, 2009).
6. Eckhouse, S., Lewison, G. & Sullivan, R. *Mol. Oncol.* **2**, 20–32 (2008).
7. Eckhouse, S. & Sullivan, R. *PLoS Med.* **3**, e267 (2006).
8. DiMasi, J. & Grabowski, H. *J. Clin. Oncol.* **25**, 209–216 (2007).
9. Reichert, J. & Wenger, J. *Drug Discov. Today* **13**, 30–37 (2008).
10. Teutsch, S., Berger, M. & Weinstein, M. *Health Aff.* **24**, 128–132 (2005).
11. Gafni, A. & Birch, S. *Soc. Sci. Med.* **62**, 2091–2100 (2006).
12. Raftery, J. *J. Eval. Clin. Pract.* **5**, 361–366 (2001).
13. Nord, E. *Med. Decis. Making* **15**, 201–208 (1995).
14. Cohen, J. *J. Med. Ethics* **22**, 267–272 (1996).
15. Tsuchiya, A. *Health Econ.* **9**, 57–68 (2000).
16. Clement, F., Harris, A. & Jing Li, J. *J. Am. Med. Assoc.* **302**, 1437–1443 (2009).
17. Mason, A. & Drummond, M. *Eur. J. Cancer* **45**, 1188–1192 (2009).
18. Tangka, F.K. *et al. Cancer* **116**, 3477–3484 (2010).
19. Jonsson, B. & Wilking, N. *Ann. Oncol.* **18** Suppl 3, iii2–iii7 (2007).
20. Verdecchia, A. *et al. Eur. J. Public Health* **18**, 527–532 (2008).
21. Gatta, G. *et al. Cancer* **89**, 893–900 (2000).
22. Verdecchia, A. *et al. Lancet Oncol.* **8**, 784–796 (2007).
23. Schiller, J. *Clin. Cancer Res.* **11**, 5030s–5032s (2005).
24. <<http://info.cancerresearchuk.org/news/archive/cancernews/2009-11-03-Lung-cancer-experts-call-for-improvements-in-NHS-care>>
25. Cohen, J., Cairnes, C., Paquette, C. & Faden, L. *J. App. Health Econ. Health Pol.* **5**, 177–187 (2006).
26. Cohen, J. & Wilson, A. *mAbs* **1**, 56–66 (2009).
27. Anonymous. NICE Guidance (National Health Service, London) <<http://guidance.nice.org.uk/TA169>> (Accessed November 1, 2009).
28. Anonymous. Pfizer's sunitinib is recommended as first-line treatment for kidney cancer patients by British health agency. <<http://www.medicalnewstoday.com/articles/138010.php>> (2009).
29. Leveque, D. *Lancet Oncol.* **9**, 1102–1107 (2008).
30. Cohen, J., Wilson, A. & Faden, L. *Food Drug Law J.* **64**, 391–403 (2009).
31. Singer, P. Why we must ration health care. *New York Times Magazine* (July 19, 2009).
32. Leigh, N., Tsao, W., Zawisza, D., Nematollahi, M. & Shepherd, F. *Lung Cancer* **51**, 115–121 (2006).
33. Markman, M. *J. Oncol. Pract.* **4**, 262 (2008).
34. Mohr, P. & Tunis, S. *Pharmacoeconomics* **28**, 153–162 (2010).

What's fueling the biotech engine—2009–2010

Saurabh Aggarwal

Last year, the biologics sector managed single-digit growth in the United States, driven mainly by products indicated for oncology, diabetes and autoimmune disorders. Lurking on the horizon, though, are challenges, such as pricing, competition and follow-on molecules.

Despite a challenging economy, sales of biologics in the United States rose by 3%, albeit at a lower rate than reported in last year's survey¹. This slowdown in sales growth was mainly due to continued safety woes for erythropoietins, manufacturing lapses related to enzyme replacement therapies, as well as market saturation and crowding. In 2009, total US sales of biologics reached ~\$48 billion. The single-digit growth was the second consecutive year of a slowdown in sales growth (Fig. 1a). 2010 doesn't look much better; although biologics sales climbed by 3.3% in Q1, by Q2 they had declined by 1.8% (Fig. 1b). Despite an overall trend of slowing growth—and for some notable blockbusters, such as Aranesp (darbepoetin alfa), even decreasing sales—a small number of products have continued to increase in sales at high rates (8–15%; Figs. 2 and 3). Thus, the biotech industry finds itself at a crossroads: several promising products are reaching the market but patent expirations, biosimilars, US healthcare reform and competition pose long-term challenges.

In the following article, I analyze the market trends observed in 2009 (and the first half of 2010) for nine classes of biologics. For each of these classes, I provide a discussion on sales volume, pricing, indication expansions, competition within biologics and from small-molecule drugs, safety issues and promising new candidates. On the basis of the sales trends, I have categorized products within each respective therapeutic class as market leaders (where sales are greatest for an indication), rising stars (where sales show rapid growth) or laggards (where sales underperformed). The methodology applied in this article is essentially the same as I used in previous biotech market overviews

Saurabh Aggarwal is a healthcare consultant based in New York.
e-mail: saurabhaggarwal@gmail.com

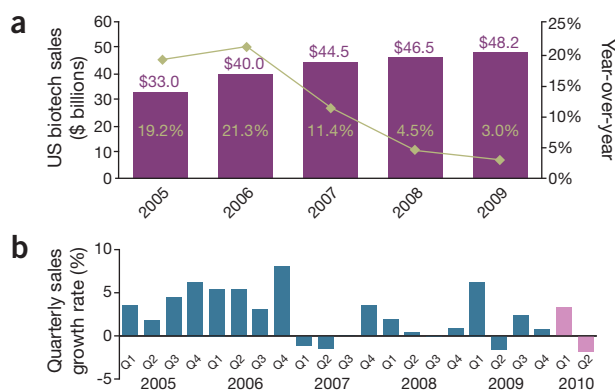


Figure 1 Growth trends in the US biotech market for biologic drugs (2005–2009). (a) Total sales and growth rate trends. (b) Quarterly sales growth for biologic drugs (2005–Q2, 2010).

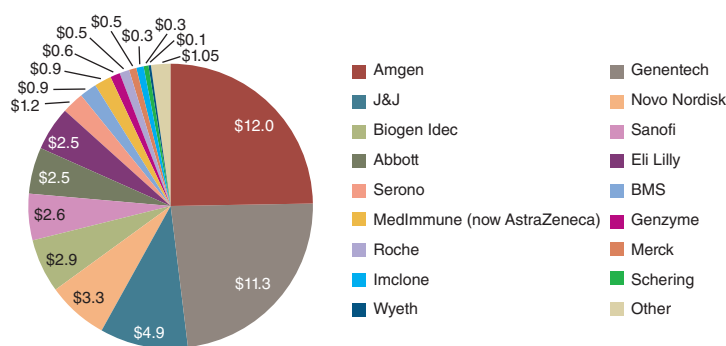
published in this journal^{1–3}. The following sections cover each of the nine classes of biologics in descending order of the sales they accrue.

Monoclonal antibodies

In 2009, monoclonal antibodies (mAbs) maintained their ranking as the best-selling class of biologics, with their US sales reaching ~\$16.9 billion—an 8.3% growth over their 2008 sales. Whereas mAbs remain one of the two most rapidly growing classes of biologics, their sales growth has slowed from last year's double-digit rate (Fig. 3). This decline in sales growth was due to saturation of the major markets as well as competition from small molecules. New products launched in 2008–2009 are showing some signs of growth, which could potentially fuel future growth of mAbs. Currently, 31 mAbs have been registered for marketing with the US Food and Drug Administration (FDA)—Pfizer (New York) withdrew its antibody conjugate Mylotarg (gemtuzumab ozogamicin) in June after a post-approval trial raised concerns about the drug's safety and efficacy. Sales of

the remaining 31 mAbs constitute ~35% of the total biologics market. The two best-selling indications for mAbs continue to be cancer and inflammatory disorders, garnering ~40% and 38% of total mAb sales in all indications, respectively. The mAbs for cancer continue to grow at a high double-digit rate (~15%), driven by new product launches, indication expansion and high unmet need⁴. The growth of anti-inflammatory mAbs (the majority of sales belonging to anti-tumor necrosis factors (TNFs)) slowed from high double-digits to mid-single digits, leading to market share wars among new and established players.

During the past year, only two blockbuster mAbs achieved sales that grew by high double-digit rates: Roche/Genentech's (Basel) Avastin (bevacizumab) against vascular endothelial growth factor (VEGF) and Abbott's (Deerfield, IL, USA) Humira (adalimumab) against TNF- α (Fig. 4). The combined sales of these two mega blockbusters reached \$5.5 billion in 2009, constituting 33% of the total mAb sales and ~60% of this sector's sales growth.



	2008 US sales (\$ billions)	2009 US sales (\$ billions)	2008 growth (%)	2009 growth (%)
Amgen	13.0	12.0	-7.2	-7.6
Genentech	10.0	11.3	10.0	12.9
J&J	4.9	4.9	-5.5	-0.7
Novo Nordisk	2.8	3.3	21.9	17.4
Biogen Idec	2.7	2.9	12.1	6.4
Sanofi	2.2	2.6	30.0	21.5
Abbott	2.2	2.5	34.8	15.3
Lilly	2.3	2.5	12.8	10.0
Serono	1.0	1.2	18.0	13.7
BMS	0.8	0.9	29.1	10.2

Figure 2 Top companies that comprise the majority of sales of biologic drugs in 2009. The pie chart shows the fraction of total biotech sales of the top 17 companies. The table shows the annual growth rates of the top ten companies. Red boxes indicate companies that had biologics sales growth >15%. For the purpose of this analysis, Rituxan US sales have been split equally between Genentech and Biogen Idec; Erbitux US sales were split 40/60 between Imclone and BMS.

Since its FDA approval in 2004 for colon cancer, Avastin's label has expanded to five more types of cancers—glioblastoma multiforme, advanced renal cell carcinoma, metastatic colorectal cancer, metastatic breast cancer and non-small cell lung cancer.

In July, however, Avastin suffered a setback when the FDA's Oncologic Drugs Advisory Committee (ODAC) recommended against its full approval for first-line treatment of metastatic breast cancer. In 2008, Genentech had received accelerated approval for Avastin's use in that indication based on results of the E2100 trial, which demonstrated a 5.5-month benefit in progression-free survival (PFS) but not overall survival (OS), compared with background paclitaxel (Taxol)^{5,6}. Two trials completed after Avastin's accelerated approval (AVADO and RIBBON-1) confirmed the PFS benefit, but they did not show any improvement in OS⁶. This ODAC recommendation has major implications for cancer drug developers because after the 2008 FDA approval of Avastin based on PFS data, some manufacturers changed the primary endpoint of their trials from OS to PFS⁷. Because 20–30% of Avastin's sales belong to metastatic breast cancer, withdrawal of this indication could slow down or lower Avastin's

sales in the near future⁶. Potential catalysts for Avastin are indication expansions for ovarian and prostate cancers.

For the past five years, Humira has been one of the most rapidly growing biologics in the United States. In 2009, Humira's sales grew by ~15%, surpassing the TNF- α market growth rate (~3%) by more than fivefold. The slowdown in the overall growth of the TNF- α market has limited Humira's growth potential to share gains in rheumatology indications and expanding use in Crohn's and psoriasis markets. Several new entrants, such as UCB's (Brussels) Cimzia (certolizumab pegal), Johnson & Johnson (J&J) Ortho Biotech's (Bridgewater, NJ, USA) Stelara (ustekinumab) and Centocor (Malvern, PA, USA)/Ortho Biotech's Simponi (golimumab), as well as follow-on biologics, pose long-term threats to Humira. A potential catalyst for Humira is indication expansion in ulcerative colitis.

In 2009, among the blockbuster mAbs, the anti-VEGF-A humanized IgG1- κ antibody fragment Lucentis (ranimizumab) grew by 20%, Remicade (infliximab), Rituxan (rituximab) and Herceptin (trastuzumab) experienced modest growth of 4–5%, and Erbitux (cetuximab) and Synagis (palivizumab)

declined by ~6–10%. The sharp increase in Lucentis sales was due to a change in a ruling by the US Centers of Medicare and Medicaid Services (CMS), as it dropped the reimbursement code for Avastin's off-label use for wet acute macular degeneration. Remicade's sales stabilization is driven by a slowdown in the anti-TNF α market, loss of market share to other products, including some cannibalization by J&J's two new anti-inflammatory mAbs, Stelara and Simponi. Herceptin and Rituxan sales have been steady for the past few years owing to saturation in drugs for lymphoma and HER2-positive breast cancer, respectively. (The FDA's recent approval of Herceptin for stomach cancer might reverse this trend in the coming year.) Erbitux's sales decline is likely due to increased use of KRAS mutation testing, which has limited its patient population to almost half of the previously broad indication in metastatic colorectal cancer.

Sales of Synagis, a humanized mAb targeted to the F protein of respiratory syncytial virus, declined last year after the American Academy of Pediatrics Committee on Infectious Disease published new guidelines, which recommended lower usage in premature infants 32 to 35 weeks' gestational age, mainly driven by the high cost of treatment and economic analyses, which did not demonstrate overall cost savings for healthcare providers⁸. This is a notable event for all manufacturers because in the past physician guidelines have rarely criticized cost of treatment. AstraZeneca (London)/MedImmune (Gaithersburg, MD, USA), the manufacturer of Synagis, suffered another setback in 2010, when an FDA advisory committee voted against approval of its follow-on version called motizumab, citing safety concerns and lack of noninferiority data⁹.

Three new mAbs on the market in 2010 were Roche/Chugai's Actemra (tocilizumab), GlaxoSmithKline's (GSK; Brentford, UK) Arzerra (ofatumumab) and Amgen's (Thousand Oaks, CA, USA) Prolia (denosumab). Actemra is a first-in-class humanized mAb that inhibits interleukin (IL)-6 receptor, indicated for moderately to severely active rheumatoid arthritis. Arzerra is a human anti-CD20, indicated for the treatment of individuals with chronic lymphocytic leukemia refractory to Fludara (fludara-bine) and Campath-1H (alemtuzumab). Prolia is a twice-a-year injectable fully human mAb that inhibits RANK ligand, indicated for treatment of osteoporosis in patients at high risk of fracture. Prolia is also in late stages of development for treatment of bone metastasis.

Several mAbs currently in late-stage clinical trials also show promise as revenue-generating franchises. Genentech/Roche's two second-generation versions of Herceptin,

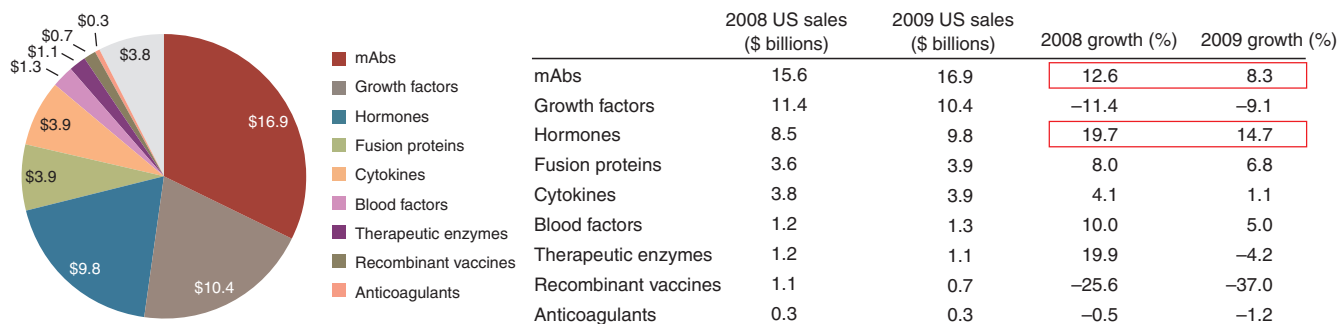


Figure 3 Top nine categories of biologic drugs in terms of US sales in 2009. The pie chart shows US sales of these drug categories. The table shows the growth rates of the categories between 2008 and 2009. The red boxes indicate the major categories showing the fastest growth rate during that period. For therapeutic enzymes, their manufacturers do not break out the US sales, so their sales were estimated assuming 20–30% of worldwide sales were generated in the United States.

trastuzumab-emtansine (Herceptin conjugated with ImmunoGen’s (Waltham, MA, USA) antimetabolic maytansine derivative DM1) and Omnitarg (pertuzumab, a humanized mAb that inhibits dimerization between HER2 and other members of the ErbB family, including epidermal growth factor receptor) have had mixed fortunes. Although trastuzumab-emtansine looked on course for success after the company filed a biologic license application (BLA) in July seeking accelerated approval on the basis of positive results from a phase 2 trial, one month later the FDA issued a ‘refuse to file’ letter stating that the trial did not meet the requirements for accelerated approval; a BLA is now planned for 2012. In the meantime, the company aims to file a BLA for Omnitarg for first-line use in HER2-positive, metastatic breast cancer sometime next year

In July, after a multicenter phase 3 trial in which BenLysta (belimumab), a human mAb targeting B-lymphocyte stimulator, met its primary endpoints, Human Genome Sciences (Rockville, MD, USA) and GSK filed for FDA approval in systemic lupus erythematosus (SLE); one month later, the FDA granted ‘priority review’ designation, and the Prescription Drug User Fee Act (PDUFA) date—the deadline for the FDA to finish the review—is now set for December 9.

In June, Bristol-Myers Squibb (BMS; Princeton, NJ) submitted a BLA for the human mAb ipilimumab against cytotoxic T-lymphocyte antigen 4 (CTLA4) in metastatic melanoma. This has also been accepted for priority review, with the PDUFA date expected to be December 25. BMS is also developing with Merck (Whitehouse Station, NJ, USA) MK 3415A, a human mAb generated using

Medarex’s UltiMab technology that targets the *Clostridium difficile* enterotoxins A and B; Merck is currently preparing to initiate a phase 3 trial of the mAb for the treatment of *C. difficile* infections and diarrhea¹⁰.

Growth factors

During the past 12 months, the growth factor sector experienced several major events, which have led to continued downward pressure on its sales. The sector—once the largest-selling category of biologics, including erythropoietin-stimulating factors (ESAs) and granulocyte macrophage colony-stimulating factors (GM-CSFs)—suffered a 9% decline in 2009, as its sales fell from ~\$11.4 billion in 2008 to ~\$10.4 billion. Compared with the growth factor sector’s peak sales in 2006, last year’s sales represent ~35% decrease. These sales belong to 11 brands of growth factors (ten original molecules) currently available in the United States, most falling into the category of ESAs and CSFs. Growth factor sales declines were due to a 12% and a 4% decrease in sales of ESAs and CSFs, respectively.

ESAs, which are sold in the United States under three brand names, Epogen (Amgen), Procrit (J&J’s Ortho Biotech) and Aranesp (Amgen) (Fig. 5), have experienced declining sales due to label changes enforced by the FDA in 2008, which restricted their use in individuals who are expected to be cured of cancer and those who have hemoglobin levels greater than 10 g per deciliter of blood. Furthermore, in 2010, four additional events occurred that are likely to depress sales of ESAs for the near term. First, FDA approved a risk evaluation and mitigation strategy to ensure the safe use of ESAs in cancer patients¹¹. Second, CMS and FDA convened two advisory meetings to discuss use of ESAs in end-stage renal disease patients, in which the committee recommended low or no use of ESAs in pre-dialysis patients and

© 2010 Nature America, Inc. All rights reserved.

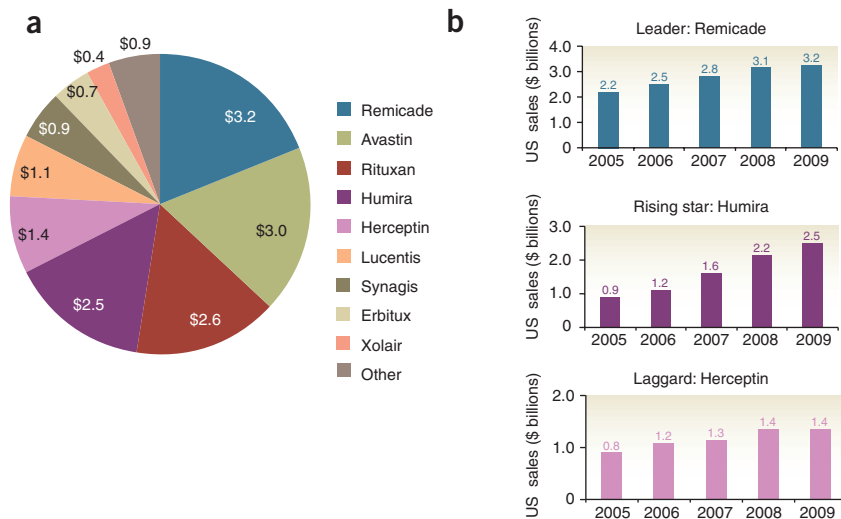


Figure 4 Trends in US sales of mAbs. (a) 2009 sales for US markets for mAbs (\$ billions). ‘Other’ includes all mAbs with sales <\$300 million per year. (b) Trends in US sales show Remicade leading, Humira rising and Herceptin lagging.

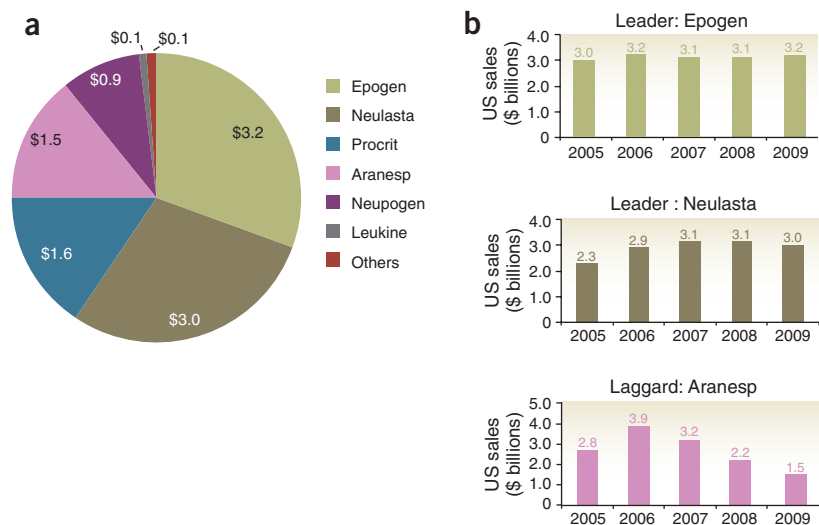


Figure 5 Trends in US sales of growth factors. (a) 2009 sales in US market for growth factors (\$ billions). (b) Trends in US sales show Neulasta and Epogen leading with Aranesp lagging.

moderate use in dialysis patients. Third, in an article published in the *New England Journal of Medicine*¹², FDA raised concerns regarding the use of ESAs for raising hemoglobin in renal patients. Fourth, in July, CMS released its final reimbursement rule that would bundle payment for ESAs for use in renal setting, which fixes the amount that is paid out, starting next January¹³. On the basis of these events, in 2010–2012, the ESA market could further decline by 20–30%, which would be a 40–45% fall compared with its 2006 peak sales.

Just a few years ago, ESAs were considered one of the most lucrative products for biosimilar/follow-on biologic development, attracting investment from small biotech and large pharmaceutical companies. However, the setbacks in terms of safety concerns for these products have significantly altered the market outlook for ESAs, making it likely that some players would rethink their plans to develop biosimilar/follow-on versions. For example, early this year, Merck announced that due to FDA's requirement for additional safety data, they discontinued the development of MK-2578, a pegylated form of erythropoietin similar to Amgen's Aranesp. Similarly in 2008, Takeda (Tokyo) and Affymax (Palo Alto, CA, USA) announced that were suspending trials for Hematide (a synthetic pegylated peptide-based ESA) in chemotherapy-induced anemia and would focus on treatment of renal disease-related anemia.

Approximately 40% of growth factor sales are derived from Amgen's two brands of CSFs, Neulasta (pegfilgrastim; pegylated human recombinant granulocyte (G-CSF)) and Neupogen (filgrastim; G-CSF), whose main indi-

cation is for use in stimulating the production of neutrophils and leukocytes in treatments related to chemotherapy and bone marrow transplants. Sales of these two products are slowly declining, mainly because of market saturation and the launch of new chemotherapies that have a lower incidence of neutropenia.

Hormones

In 2009, the only class of biologics that experienced double-digit growth in sales was that of hormones. Last year hormone sales reached ~\$10 billion at an overall growth rate of 15% over last year (Fig. 3). In 2010, hormones will likely replace growth factors as the second-best-selling class of biologics, assuming similar growth rates and continued slowdown in sales of ESA and CSF products. The rapid sales growth of hormones have been driven mainly by high double-digit growth in insulin analogs.

Last year, US sales of insulin products reached \$7.2 billion, constituting 76% of the hormone sector sales, with 21% growth over their 2008 sales. This increase in sales was primarily due to large increases in sales of rapid-acting and long-acting analogs (Fig. 6). Among rapid-acting insulins, Novo Nordisk's (Bagsvaerd, Denmark) Novolog (insulin aspart) and Eli Lilly's (Indianapolis) Humalog (insulin lispro) product revenues grew by 26% and 19%, respectively. Despite a few safety events reported in 2009, Sanofi-aventis's (Paris) long-acting insulin Lantus (insulin glargine) sustained its high growth rate (22% in 2009 compared with 30% in 2008), gaining the status of most rapidly growing mega blockbuster biologic (second in ranking is Humira).

A new entrant to the hormones market is Novo Nordisk's once-daily glucagon-like peptide-1 (GLP-1) receptor agonist Victoza (liraglutide). This is the second GLP-1 analog to hit the market after Amylin's (San Diego) twice-daily administered Byetta (exenatide), sales of which declined in 2009, mainly due to its linkage to a dangerous form of pancreatitis¹⁴. In clinical trials, GLP-1 analogs have been shown to improve glycemic control in adults with type 2 diabetes, with an additional benefit of weight loss in some people. Two safety concerns have been raised about these drugs, however, one being the risk of medullary thyroid cancer, which was observed in mice and rats, and the second being pancreatitis, which was observed in the post-marketing reports on Byetta submitted to the FDA and in clinical studies on Victoza¹⁵. Based on these risks, FDA has asked manufacturers of GLP-1 analogs to conduct additional animal studies and to maintain post-marketing registries¹⁵. Further bad news for this class of drugs came in October, when the FDA sent Amylin a complete response letter for its once-weekly version, bydureon.

Fusion proteins

There are six fusion proteins currently available on the US market. In 2009, the combined US sales of these six brands were \$3.8 billion, a 1% increase over their 2008 sales (Fig. 3). More than 80% of these sales belong to Amgen's Enbrel (recombinant TNF- α receptor fused to IgG fragment), whose unit sales have been declining in recent years. Although Enbrel continues to be the best-selling biologic in the United States, it is experiencing a declining trend in its sales. Because Enbrel's wholesale price was increased by 7–8% in 2009 and 5% in the second quarter of 2010, on a per-unit basis, its sales actually declined by mid single-digit rates. Enbrel's share in rheumatology declined from 36% to 33%, because of continued share gains by Humira. Interestingly, in the dermatology segment Enbrel also lost significant market share (its total share declined from 62% to 49%) to Stelara, a relatively new player, whose market share is now 13% in this segment. As is evident from Stelara's share gains, non-anti-TNF products are now competing for market share with established anti-TNF players.

Launched in 2006, BMS's Ocrencia (abatcept), an IgG fused to the extracellular domain of CTLA4) is one of the most rapidly growing biologics in this category. It is indicated for treatment of individuals with rheumatoid arthritis who have had an inadequate response to other drugs, primarily anti-TNF- α mAbs. Because rheumatoid arthritis is a chronic and high-incidence disease, a substantial number of people do not respond, stop responding or

develop antibodies to anti-TNFs, becoming eligible for Orencia, which works by blocking CTLA4 signaling, a new mechanism of action for treating arthritis¹⁶. During the past four years, its sales have climbed to \$480 million. Because anti-TNF- α molecules are used in several other autoimmune indications, such as Crohn's, ulcerative colitis and psoriasis, patients who do not respond to anti-TNF- α drugs could potentially become eligible to receive Orencia. BMS is conducting phase 3 clinical trial for subcutaneous formulation and is further testing this drug in psoriatic arthritis and lupus nephritis patients.

A likely new entrant in this sector is belatacept, a follow-on version of Orencia. Early this year, a FDA advisory committee had recommended approval for this product for kidney transplant patients. However, in a complete response letter, the FDA has asked BMS for 36-month data from the ongoing phase 3 trials and the proposed risk evaluation and mitigation strategy. This is probably related to the incidence of progressive multifocal leukoencephalopathy (PML) and T-cell lymphoma, which were observed in ongoing phase 3 clinical trials. Because belatacept differs from Orencia by only two amino acids¹⁷, this product provides a potential case study for developers of biosimilars.

Sales of Amgen's Nplate (romiplostim), a 60 kDa peptide with a thrombopoietin receptor-binding domain that is indicated for long-term treatment of adult chronic immune thrombocytopenic purpura, more than doubled from ~\$40 million in 2008 to \$85 million in 2009. Another fusion product that is far from a blockbuster is Regeneron's (Tarrytown, NY, USA) Arcalyst (IL-1 trap); indeed, Arcalyst is a niche player, as cryopyrin-associated periodic syndromes affect only a few hundred people in the United States. As per company reports, US sales of the drug doubled from \$11 million in 2008 to ~\$22 million in 2009. Regeneron is testing Arcalyst for treatment of gout disorders, which is a much broader indication and could potentially lead to larger sales.

Cytokines

In 2009, US sales of cytokines grew to \$3.8 billion, largely driven by steep price increases. Four kinds of cytokines are currently available: interferon (IFN)- α , IFN- β , IFN- γ and the interleukins. IFN- β brands, indicated for multiple sclerosis (MS), are the biggest sellers in the cytokines markets, constituting 76% of the total sales of this sector (Fig. 7a).

Last year, sales of IFN- β rose by 7%, driven by 18–27% price increases for all three marketed brands. As previously reported, all

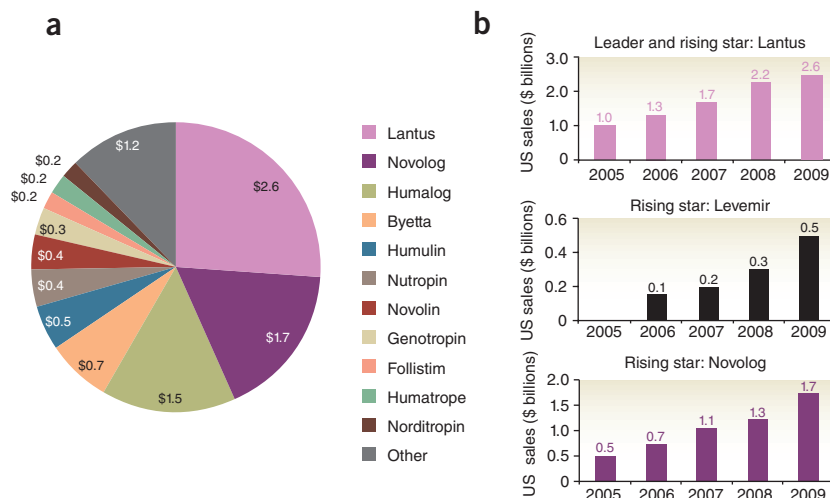


Figure 6 Trends in US sales of recombinant hormones. (a) 2009 sales in US market for recombinant hormones (\$ billions). (b) Trends in US sales show Lantus to be the leader, Novolog a rising star and Levemir new on the block. 'Other' includes all hormones with sales of <\$200 million per year.

three manufacturers raised the prices of their drugs a number of times because of saturation of the MS market. In the future, the unit sales of interferon- β in MS are likely to fall, mainly owing to competition from a small-molecule drug called Gilinea (fingolimod, a sphingosine-1-phosphate receptor modulator; Novartis, Basel), which was approved by the FDA in September. In a head-to-head trial called TRANSFORMS, Gilinea demonstrated much higher efficacy than IFN- β —showing a 38–52% reduction in relapse rate¹⁸.

IFN- α brands, indicated for hepatitis C, generated US sales of \$800 million in 2009, a 3% decline from their 2008 sales, mainly because some physicians are putting patients on hold, awaiting two potential breakthrough pipeline small-molecule products—telaprevir (Vertex, Cambridge, MA, USA) and boceprevir (Merck), which are currently in late stages of clinical development. These products have been shown to have 28–31% higher sustained virologic response than the standard of care, which is pegylated IFN- α and ribavirin (Rebetol, Copegus, Virazole)^{19,20}. Add-on treatment with small molecules in development could make many more patients eligible for treatment as it is estimated that there are ~3 million chronic hepatitis C patients, of which 75% remain unaware of their infection²¹. As more patients get treated with triple combination therapy, it could create more need for pegylated interferon.

Therapeutic enzymes

Therapeutic enzymes have been a relatively small but rapidly growing sector of the biological market. Until 2008, this sector experienced

high double-digit growth. In 2009, however, eight different brands of enzymes generated US sales of ~\$1.05–1.11 billion, a 5–10% decrease over their 2008 sales (Figs. 3 and 7b). This decline in sales was due to manufacturing lapses faced by Genzyme (Cambridge, MA, USA) for Cerezyme (imiglucerase) and Fabrazyme (agalsidase- β), which led to serious shortages of these two products. For most therapeutic enzymes, their manufacturers do not break out the US sales, so, based on previous years' trends, I have estimated that 25–30% of worldwide sales were generated in the United States.

The manufacturing problems Genzyme faced at its Allston, Massachusetts, plant most severely affected the supply of its blockbuster product Cerezyme (human recombinant glucocerebrosidase), whose sales declined by ~35% in 2009. Because Cerezyme was the only FDA-approved therapy indicated for patients with Gaucher disease, FDA granted treatment protocol to Protalix (Carmiel, Israel) and Shire (Dublin) for their biologic glucocerebrosidase replacement candidates Vpriv (velaglucerase alfa) and taliglucerase alfa, respectively. These events have provided Shire and Protalix with an easy entry point into the Gaucher disease market. With recent FDA approval of Shire's Vpriv and a PDUFA date of early 2011 for Protalix's taliglucerase alfa, this market will have three players in 2011. Recently, Shire announced that it has priced Vpriv at \$1,350 per 400 unit vial, a ~15% discount on Cerezyme, which is an early sign of price wars in this ultra-orphan indication of only 1,500 patients in the United States.

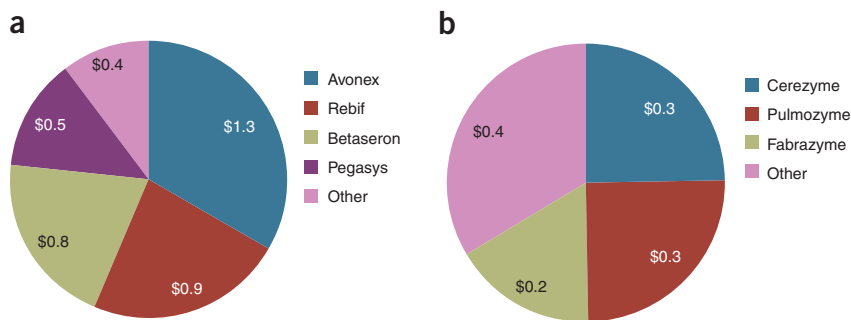


Figure 7 US sales of cytokines and therapeutic enzymes (\$ billions). (a) Cytokines sales, showing the top four brands. (b) Therapeutic enzyme sales, showing the top three brands. ‘Other’ includes all drugs with sales of <\$200 million per year.

Recombinant vaccines

During 2009, the recombinant vaccines sector experienced a sharp fall, from \$1 billion to \$0.7 billion (Figs. 3 and 8a), overall, a 37% decrease over last year. Unlike products indicated for chronic diseases, vaccines are typically prescribed one or a few times (depending upon the need for booster shots) during an individual’s lifetime. This leads to a relatively large uptake 1–3 years after launch, followed by a decline to a stable and lower level of sales. More than 70% of US sales of vaccines are attributed to Merck’s human papilloma virus (HPV) vaccine Gardasil, indicated for prevention of HPV (types 6, 11, 16 and 18) in girls and women between the ages of 9 and 26 years and for prevention of genital warts caused by HPV (types 6 and 11) in boys and men between the ages of 9 and 26 years. Merck won FDA approval for the male indication in October 2009. According to Merck’s recent quarterly report, the male indication is not currently a large contributor of sales. Although this new indication could present an important opportunity, one challenge in penetrating this market is the reported low cost effectiveness of vaccinating males with Gardasil^{22,23}. Cost effectiveness is not yet formally used by US physicians or payers, but it has increas-

ingly surfaced in debates on Gardasil, perhaps because Gardasil was priced aggressively when it was first marketed. Last year, the US Centers for Disease Control and Prevention’s Advisory Committee on Immunization Practices voted to permit, but not recommend, routine vaccination of boys, at least partly based on the vaccine’s low cost-effectiveness in boys²³. Meanwhile, GSK’s Cavarix, approved in 2009, is the second HPV vaccine to make it to market. As a second entrant to the market, GSK is offering this vaccine at a discount compared to Gardasil.

Blood factors

Blood factors represent one of the smaller markets in the biologics sector; the seven brands that comprise this category generated US sales of ~\$1.28 billion in 2009. Their sales grew by 5% in 2009 (Figs. 3 and 8b). Three blood factors make up this market: factor VIII, factor VIIa and factor IX. More than half of the sales in this category belong to Novo Nordisk’s NovoSeven (eptacog alfa; recombinant activated factor VII). In 2009, sales of NovoSeven RT (eptacog alfa formulated with sucrose and L-methionine), which is stable at room temperature, more than doubled to ~\$0.5 billion, replacing the sales of its older version called NovoSeven. The com-

bined sales of factor VIII and factor IX brands were ~\$765 million in 2009 in the United States, an 8% increase over their 2008 sales.

Anticoagulants

The thrombolytics/anticoagulants segment is one of the smaller biologics markets and has been experiencing consistently declining sales. The past five years have seen combined US sales for molecules in this category contract from \$347 million in 2004 to \$328 million in 2009 (Figs. 3 and 8c). This decline in sales is likely due to several new small-molecule anti-coagulants that have reached the market.

Conclusions

Last year marked the second year where sales from the biologics sector slowed from their historical high double-digit rate to mid-single digits. This slowdown is due to a combination of factors—safety issues, manufacturing lapses, saturation of approved indications, crowding and the ongoing US recession. The safety issues of ESAs (historically, the best-selling category of biologics) has led to a ~37% decrease from their 2006 peak sales. Several events in 2010, such as the FDA’s risk evaluation and mitigation strategy, CMS Advisory meeting and the bundling of payments, all suggest continued decline in sales of ESAs, likely eroding a few more percentage points from the near-term growth of the overall sector. Though a one-time event, Genzyme’s manufacturing problems also led to a 35% decline in sales of its ultra-orphan, billion dollar product. At the same time, several blockbuster biologics, such as Rituxan, Herceptin, Enbrel and Gardasil, have experienced a slowdown or decline in their sales because of high penetration in their FDA-approved indications and/or competition from other biologics. Although sales of medicines indicated for very severe diseases are perceived to be resistant to the performance of the overall economy, there are increasing signs that due to the ongoing recession more

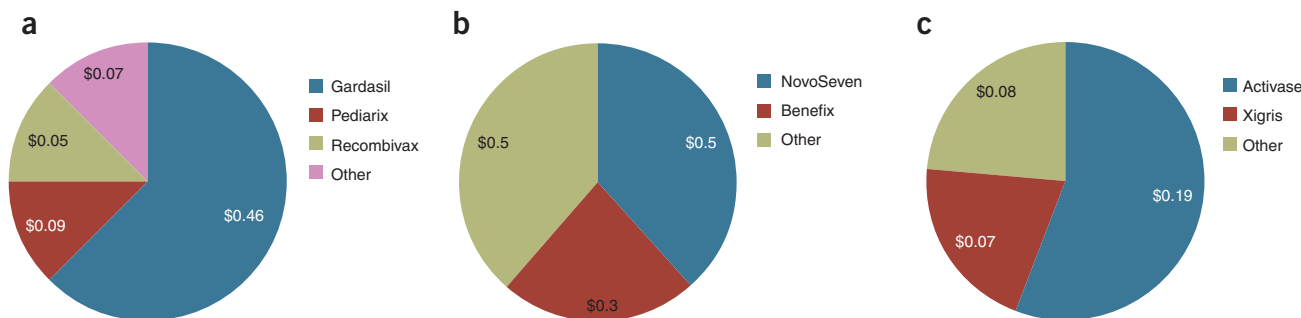


Figure 8 US sales of recombinant vaccines, blood factors and anticoagulants in (\$ billions). (a) Recombinant vaccines, dominated by one brand. (b) Blood factors dominated by two brands. (c) Recombinant anticoagulants market, which is dominated by two brands. ‘Other’ includes all drugs with sales of <\$50 million per year.

Americans are passing on healthcare products and services²⁴.

The outlook for the biologics market is mixed. One near-term threat to sales of this sector is the looming patent expiration for several blockbuster biologics brands, whose sales could slow down markedly if lower priced follow-on biologics are approved in the US market. With the recent passage of the US healthcare reform legislation and subsequent establishment of the official pathway for FDA approval of follow-ons, it is likely that several such products could be launched during the next 2–3 years. The recent price wars seen in some classes of biologics—human growth hormone, ultra-orphan enzyme replacement products and vaccines—are signs of imminent pricing pressure for products targeted for follow-on development. The new 23.1% rebate for US Medicaid programs and discounts under the 340B provision are also likely to affect the price and sales of some biologics.

A longer-term threat to sales of biologics is new payment reform, such as the recently passed CMS bundled payment system for dialysis centers. With fully bundled reimbursement mechanisms, healthcare providers would be sensitive to the cost of therapies, creating an incentive to use lower price products.

On a positive note, several recently launched products and other late-stage pipeline products look promising and could fuel further growth in this sector. Amgen's Prolia (FDA approved

for osteoporosis, BLA filed for bone metastases) is expected to be a blockbuster product. With the approval of Dendreon's (Seattle) Provenge (sipuleucel-T) prostate cancer vaccine, the first allogeneic cell vaccine approved in the United States, a new type of biologics might also be accruing multimillion dollar sales by the time I survey the sector again. Biologics in late stages of development—BMS's belatacept (kidney transplant) and ipilimumab (melanoma), Roche/Genentech and Immunogen's trastuzumab-DM1 (breast cancer), Human Genome Sciences/GSK's belimumab (lupus) and Regeneron/Sanofi's Arcalyst (gout)—all are expected to be multibillion dollar products. Additionally, on-the-market mAbs and insulin analogs are likely to experience continued high growth rates in the near term.

1. Aggarwal, S. *Nat. Biotechnol.* **27**, 987–993 (2009).
2. Aggarwal, S. *Nat. Biotechnol.* **26**, 1227–1233 (2008).
3. Aggarwal, S. *Nat. Biotechnol.* **25**, 1097–1104 (2007).
4. Aggarwal, S. *Nat. Rev. Drug Discov.* **9**, 427–428 (2010).
5. Pollack, A. FDA extends Avastin's use to breast cancer. *New York Times* 23 February 2008 <<http://www.nytimes.com/2008/02/23/business/23drug.html>>
6. Peterson, M. & Doherty, D. Roche may lose \$1 billion a year on Avastin change. *Bloomberg Businessweek* 21 July 2010 <<http://www.businessweek.com/news/2010-07-21/roche-may-lose-1-billion-a-year-on-avastin-change.html>>
7. Anonymous. Pink sheet—FDA will revisit appropriate use of PFS endpoints at advisory committee. *Friends of Cancer Research Newsletter*, 2008 <<http://www.focr.org/pink-sheet-fda-will-revisit-appropriate-use-of-pfs-endpoints-at-advisory-committee.html>>
8. Krilov, L.R. *et al. Pediatrics* **124**, 1682–1684 (2009).
9. Lowry, F. FDA panel nixes licensing request for motavizumab. *Medscape Medical News* (2010). <<http://www.medscape.com/viewarticle/722903>>; June 3, 2010.
10. Lowy, I. *et al. N. Engl. J. Med.* **362**, 197–205 (2010).
11. FDA approves a risk evaluation and mitigation strategy (REMS) to ensure the safe use of Erythropoiesis-Stimulating Agents (ESAs). *US Department of Health & Human Services, US Food and Drug Administration* 16 February 2010 <<http://www.fda.gov/AboutFDA/CentersOffices/CDER/ucm200847.htm>>
12. Unger, E.F., Thompson, A.M., Blank, M.J. & Temple, R. *N. Engl. J. Med.* **362**, 189–192 (2010). <<http://www.cms.gov/ESRDPayment/>>
13. Information for Health care Professionals: Exantide (marketed as Byetta) 8.2008 update <<http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafety/InformationforPatientsandProviders/ucm124713.htm>>.
15. Parks, M. & Rosebraugh, C. *N. Engl. J. Med.* **362**, 774–777 (2010).
16. Genovese, M.C. *et al. N. Engl. J. Med.* **353**, 1114–1123 (2005).
17. Vincenti, F. *et al. N. Engl. J. Med.* **353**, 770–781 (2005).
18. Cohen, J.A. *et al. N. Engl. J. Med.* **362**, 402–415 (2010).
19. McHutchison, J.G. *et al. N. Engl. J. Med.* **360**, 1827–1838 (2009).
20. Anonymous. In pivotal phase 3 studies, Merck's investigational medicine boceprevir helped majority of patients with chronic Hepatitis C genotype 1 infection achieve sustained virologic response, the primary endpoint of the studies. *Merck Newsroom* 4 August 2010 <http://www.merck.com/newsroom/news-release-archive/research-and-development/2010_0804.html>
21. Colvin, H.C. & Mitchell, A.E. (eds.) *Hepatitis and Liver Cancer: A National Strategy for Prevention and Control of Hepatitis B and C* (The National Academies Press, 2010).
22. Kim, J.J. & Goldie, S.J. *BMJ* **339**, b3884 (2009).
23. Peres, J. *J. Natl. Cancer Inst.* **102**, 838–840 (2010).
24. Kelley, R.K. & Venook, A.P. *N. Engl. J. Med.* **363**, 596–598 (2010).

A shadow falls over gene patents in the United States and Europe

Gareth Morgan & Lisa A Haile

Will a pair of court decisions that restrict the protection offered to DNA-based claims reduce financial incentives and thus chill investment?

Two recent cases in the United States and the European Union (EU) have induced a state of anxiety in gene patent stakeholders—including not only holders of patents directed at gene sequences but also patent attorneys who draft and file patents for such claims. The first case, a reference¹ to the Court of Justice of the EU (the “Court”) for a preliminary ruling on the interpretation of the Biotechnology Directive², has cast doubt on the enforceability of claims directed at DNA sequences. The key issue that the Court has been asked to consider is whether claims directed at DNA sequences only protect the claimed DNA sequences when the corresponding DNA molecules are in a functional context. In the second case, Judge Robert W. Sweet provided a ruling on March 29, 2010, invalidating claims to both genes and methods of diagnosis based on the analysis of the native genes and mutations in the genes³. The decision is potentially far-reaching as it may affect the validity of many patents containing DNA-based claims in the biotech industry for which investors have invested large sums of money.

Monsanto v. Cefetra

The reference to the Court was made by the District Court of The Hague in the Netherlands in the case of *Monsanto v. Cefetra*⁴. Monsanto, lacking patent protection for its first generation Roundup Ready technology in Argentina, had sought to prevent Argentine genetically modified (GM) soy imports into a number of European territories through the assertion of its European patent rights. This patent did not have any claims to soy meal but did claim the

DNA sequence used in the genetic modification process. The case can therefore be summarized by the question, Can a patent claim to a DNA sequence protect any product containing DNA molecules of the claimed sequence?

The important section of the Biotechnology Directive on which the Court ruled is Article 9:

“The protection conferred by a patent on a product containing or consisting of genetic information shall extend to all material, save as provided in [Article 5(1)][paragraph 3(a) above], in which the product is incorporated and in which the genetic information is contained and performs its function.”

Experiments performed in the Dutch case proved the DNA sequences claimed in the Monsanto patent⁵ were present. To prove its claim, Monsanto therefore needed the Court to rule that Article 9 should be interpreted permissively and that claims to DNA sequences do protect any product containing molecules of the claimed sequence. However, with a display of linguistic gymnastics and little meaningful analysis, the Court ruled that Article 9 should be construed restrictively.

The Court ruled that the many express references to the requirement for the inclusion of the function of a gene within a patent should be read as incorporating the function of the gene into the concept of the DNA sequences as inventions and, correspondingly, claim scope in such patents should be equally restricted. Importantly, the Court also ruled that the Biotechnology Directive created a single regime of purpose-bound protection for DNA sequences within the EU—no deviation from this regime was to be permitted by individual member states no matter whether the patent in question was

filed before or after the implementation of the Biotechnology Directive. Often known in Europe as the ‘gold plating’ of directives, it is usually permissible only where a directive clearly sets out to ensure a minimum standard for a particular legislative regime.

Analysis of the ruling suggests the Court has fallen into error. Yes, the function of a gene must be disclosed within the application but this is required to satisfy the industrial applicability (or utility) requirement for patentability under the European Patent Convention⁶. The failure to appreciate this suggests a grave lack of understanding of fundamental patent law on the part of the Court. This is perhaps not surprising given this was the Court’s first-ever decision on substantive patent law. In relation to any other new and useful chemical, the disclosure of one use only of the new chemical is sufficient to give the inventor a monopoly over all uses of that chemical for the term of the patent. The Court now considers this rule should be modified for DNA-based inventions. Further, the requirement to include the functionality of the gene was introduced to prevent the mass patenting of expressed sequence tags, not to restrict the scope of properly granted patent claims directed at genes whose function had been characterized and disclosed in the patent.

Remarkably, in also holding that its ruling was not contrary to the Trade Related Aspects of Intellectual Property Rights agreement and did not unfairly prejudice the legitimate rights of patentees⁷, the Court did not attempt to address or analyze the effect that its ruling has in relation to claims to “isolated” DNA sequences or the ability of patentees to exploit their newly restricted DNA-based claims. The term ‘isolated’ finds its way into numerous European gene patents possibly because it is a drafting requirement in the United States. In both the English⁸ and

Gareth Morgan and Lisa A. Haile are at DLA Piper LLP, London, UK.
e-mail: gareth.morgan@dlapiper.com;
lisa.haile@dlapiper.com

Dutch parallel cases in the Monsanto Roundup Ready litigation, the national courts held that the term 'isolated' restricts the scope of the claim to DNA sequences that are separated out from other molecular species ready to be used in laboratory experiments⁹. Henceforth, patent claims to isolated DNA can now never be infringed in Europe as isolated DNA is, by definition, not functional.

This potential erosion of the patent protection offered to patentees of DNA-based inventions could have an impact on innovation within the sector. Further, the 'use' patents that have been proposed as the answer to any criticism that DNA-based inventions will not be given proper protection¹⁰ are notoriously difficult to enforce and have recently been criticized as not encouraging innovation to 'reproduce' old medicinal products^{11,12}.

Myriad in the United States

In 1980 the US Supreme Court¹³ upheld the first patent in the field of genetic engineering for the invention of oil-ingesting bacteria. The decision overturned the US Patent and Trademark Office (USPTO)'s pronouncement that life forms were not patentable. By holding that "anything under the sun made by man is patentable," any live, human-made microorganisms (as well as isolated nucleotide sequences) were deemed patentable subject matter under 35 USC §101.

The American Civil Liberties Union, the Public Patent Foundation and others filed a lawsuit in 2009 (ref. 3) challenging patents on two human genes associated with breast and ovarian cancer. It was the plaintiffs' contention that these gene patents stifle research that may lead to important cures and potentially restrict patient options with regard to treatments. The lawsuit, filed in the US District Court for the Southern District of New York against the USPTO and owners of the patents on mutations in the *BRCA* genes, Myriad Genetics and the University of Utah Research Foundation, has the potential to be a landmark case for the biotech industry.

The patents at the center of the lawsuit are based on the initial identification of mutations within two genes, called *BRCA1* and *BRCA2*, which appear to be the cause of the majority of cases of hereditary breast and ovarian cancers. The lawsuit states that the patents on the two human genes are unconstitutional and invalid because "human genes are products of nature, laws of nature and/or natural phenomena, and abstract ideas or basic human knowledge or thought."

In the March ruling, Judge Sweet struck down the Myriad patents, ruling that DNA merely contains information, and that DNA molecules of the same sequence, whether in the body or isolated, still contain the same information. It

followed that he found isolated DNA "is not markedly different from native DNA as it exists in nature." He therefore ruled that it did not constitute patentable subject matter under 35 USC §101. This ruling has the industry in some turmoil as gene patents have been issued in the United States for decades and they are not likely to stop being issued by the USPTO based on this decision by Judge Sweet.

Currently, thousands of patents have been issued in the United States to hundreds of universities, companies and other researchers on human DNA sequences. Judge Sweet's decision will be the subject of an appeal to the US Court of Appeals for the Federal Circuit (CAFC) and it is hoped that any discrepancies between his reasoning and previous authorities will be resolved, at least clarifying the scope and/or enforceability of such patents.

Conclusions

The Court of Justice of the EU has now established an EU-wide regime of "purpose-bound" protection for DNA-based inventions. This regime even applies to patents filed before the Biotechnology Directive was implemented in Europe. Although some member states implemented the Biotechnology Directive in this manner (most notably France and Germany), most did not. For patentees within the agricultural biotech sector, no derivative or processed products will be protected by claims to DNA sequences. A new form of claim drafting will need to be adopted to obtain protection on such products. For all other patentees within this field, it removes a broad layer of protection that may erode the value of key patents within the company's portfolio. The significant difficulties this ruling creates for the exploitation and enforcement of DNA-based claims are not addressed in the Court's ruling, despite it finding in less than half a page of text that patentees' legitimate interests were not being unfairly prejudiced.

In the United States, there is doubt as to whether, and to what extent, the CAFC will uphold Judge Sweet's decision. However, the *BRCA* patents were filed many years ago, at a time when diagnostic tests relied more on the 'one gene, one mutation' premise. The diagnostics industry today is based more on 'multiplex' testing, or testing multiple genes at once, as well as testing entire genomes. For single-nucleotide polymorphism (SNP) analyses, multiple SNPs over large fragments of DNA are typically tested because it is thought that genetic linkage, rather than a single mutation, is likely responsible for a trait. Thus, should Judge Sweet's decision be upheld at the CAFC, we may still see a stimulation of innovation in the discrete area of the diagnostics industry.

However, of broader concern is that gene patents were, and still are, the foundation of the biotech industry. Whether patents are for genes encoding therapeutic proteins such as insulin and growth hormone, antibodies such as Herceptin (trastuzumab) and Rituxan (rituximab), or mutated genes that are indicative of breast or ovarian cancer, precluding patents on this subject matter may inhibit the progress that has been made in the biotech industry over the 30 years since *Chakrabarty* was decided. There is a real danger that Judge Sweet's ruling will have this effect if upheld by the CAFC.

Generally speaking, and given the rate of recent advances, the patent system on both sides of the Atlantic is working as intended. Patentees are given a limited monopoly for their inventions and after the patent term expires, the patented invention is placed into the public domain. Although it is important not to lose sight of one of the central purposes of the patent system, which is to encourage public disclosure of new inventions, the reality is that in the biotech field, companies and researchers are motivated by patents either to improve upon old technologies or to develop new and innovative drugs or tests partly for the monopoly offered and partly to design around existing patents. Restricting the protection offered to DNA-based claims will reduce the available incentives, the consequence of which will be to reduce the investments companies are prepared to commit to such research. In the United States, much now rests on how the CAFC reacts to Judge Sweet's ruling. In the EU, the Court has left industry and legal practitioners with the problem of trying to work around bad law that unfairly discriminates against DNA-based inventions.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Case 428/08 *Monsanto Technology LLC v. Cefetra BV & Others*.
2. Directive 98/44/EC, OJ 1998 L 213, p.13.
3. Association for Molecular Pathology *et al. v. United States Patent and Trademark Office, et al.* No. 09 Civ. 4515 (S.D.N.Y.).
4. *Monsanto Technology LLC v. Cefetra BV & Others*, 249983/HA ZA 05-2885 (Hague District Court 2008).
5. European patent EP 0 546 090 Barry, G.F. Kishore, G.M. & Padgette, S.R., filed 28 August 1991.
6. European Patent Convention, as given effect in the recent case of *Eli Lilly v. HGS* [2010] EWCA Civ 33.
7. Paragraphs 70-77 of the Court's ruling in case 428/08, *Monsanto Technology LLC v. Cefetra BV & Others*.
8. *Monsanto Technology LLC v. Cargill parties*, EWHC 2257 (Pat), (10 October 2007).
9. Cohen, S. & Morgan, G. *Nat. Biotechnol.* **26**, 289-291 (2008).
10. See last sentence of paragraph 36 of the Advocate General's opinion in case 428/08, *Monsanto Technology LLC v. Cefetra BV & Others*.
11. Anonymous. *Nature* **465**, 267-268 (2010).
12. Morgan, G. More patent protection for medicines with a new purpose. *Nature* **465**, 1005 (2010).
13. *Diamond v. Chakrabarty*, 447 US 303 (1980).

Recent patent applications in biological imaging

Patent number	Description	Assignee	Inventor	Priority application date	Publication date
US 20100217023	New monomeric amidophosphine-metal complexes useful for, e.g., biological imaging or photochemical catalysis. The complexes are cost effective, have excellent emission and luminescence properties, longer lifetimes and good quantum efficiency.	Dempsey J, Miller AJ, Peters JC	Dempsey J, Miller AJ, Peters JC	7/28/2006	8/26/2010
US 20100208270	An interferometric detection metrology system that has a fiber optic splitter including an arm operably coupled to a broadband source and another arm directing light onto a sample; for use in, e.g., biological imaging.	Devaraj B, Kulkarni MD, Ma S	Devaraj B, Kulkarni MD, Ma S	2/19/2009	8/19/2010
US 7778296	An optical microcavity arrangement, e.g., a photonic crystal cavity arrangement with a colloid material layer comprising emitters that couple photon emission to cavities in a predetermined wavelength range; for use in a laser device.	Stanford University (Palo Alto, CA, USA)	Englund DR, Fushman I, Vuckovic J	5/12/2006	8/17/2010
WO 2010045421	A photoacoustic imaging method for prostate imaging involves focusing waves through the acoustic system that comprises the multicomponent acoustic lens, and imaging the focused waves in two dimensions.	University of Rochester (Rochester, NY, USA), Rochester Institute of Technology (Rochester, NY, USA)	Dogra VS, Knox WH, Rao NAHK	10/15/2008	4/22/2010, 7/29/2010
JP 4507132	An imaging device, e.g., digital camera, that has an auto focus mechanism that matches focus of the digital camera using light emitted by a luminous dummy object; for imaging in the detection and analysis of biological objects.	Liponics (Tokyo)	Ogasawara M, Tsuruta K	6/19/2009	7/21/2010
WO 2010060618	A new monomeric variant of the red fluorescent protein eqFP611 having an altered amino sequence featuring specific amino acid exchanges; useful as a labeling substance in biological and/or medical imaging.	University of Ulm (Ulm, Germany)	Kredel S, Nienhaus GU, Oswald F, Wiedenmann J	11/26/2008	6/3/2010
US 20100123155	A formulation for a light-emitting device, comprising a population of semiconductor nanoparticles incorporated into discrete microbeads having an optically transparent medium, where the medium is embedded in a host light-emitting diode encapsulation medium.	Nanoco Technologies (Manchester, UK)	Harris J, Pickett N	11/19/2008	5/20/2010
CN 101706617	A filter with acoustic-optic deflectors placed in a front focal point of a lens and a back focal length of another lens, respectively, and a liquid crystal space light modulator whose phase is obtained by specific formula; for use in a micro-second laser biological imaging field, light storage field and a micro-machining field.	Shenzen Advanced Technology Research Institute (Shenzhen, China)	Li D, Zhang C	11/18/2009	5/12/2010
WO 2010042948	A metal-organic framework building block comprising a tetratopic carboxylic acid and a metal, where the block is robust, noncatenated, 3-dimensional, and permanently microporous; for gas adsorption separation (e.g., carbon dioxide, hydrogen, nitrogen and methane from the gas mixture) and also for catalysis, gas storage, sensing, biological imaging and drug delivery.	Northwestern University (Evanston, IL, USA), Farha OK, Hupp JT	Farha OK, Hupp JT	10/10/2008	4/15/2010
US 20100094134	A medical imaging method for diagnosing, e.g., tumors in breast, involving scanning a tissue volume including a biological entity to quantitatively reconstruct the functional parameters of the biological entity from structural parameters.	University of Connecticut (Farmington, CT, USA)	Gamelin J, Zhu Q	10/14/2008	4/15/2010
WO 2010005550	A caged semiconductor nanocrystal, that is, quantum dot, comprising semiconductor nanocrystal and chemical moiety bound to or associated with the semiconductor nanocrystal; for use as a super resolution probe used in the detection of detectable substances in organic or biological materials.	Regents of the University of California (Oakland, CA, USA)	Cohen B	7/7/2008	1/14/2010

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>).



Plant natural products from cultured multipotent cells

Susan Roberts & Martin Kolewe

Cultured cambial meristematic cells could enable large-scale production of certain natural products.

Natural products from plants have an important place in the history of pharmaceuticals as a rich source of leads and drugs¹. Yet research into natural products has declined steadily over the past two decades, eclipsed by high-throughput screening and combinatorial chemistry and by uncertainties regarding the feasibility of large-scale manufacture². In this issue, Lee *et al.*³ show that these concerns might be overcome by culturing cambial meristematic cells (CMCs), multipotent plant cells that give rise to the vascular tissues xylem and phloem. From a commercial perspective, CMCs could supersede existing plant cell culture methods for generating natural products.

Owing to their structural complexity, most natural products cannot be produced on an industrial scale by chemical synthesis. Extraction from plants is often not feasible as the plants can be rare or slow growing. These supply issues came to national attention in the United States during development of the anticancer drug paclitaxel (Taxol) in the early 1990s, when legislation was enacted to protect the pacific yew (*Taxus brevifolia*) from overharvesting and to mandate exploration of alternative sources of paclitaxel. One solution that emerged was culture of dedifferentiated cells (DDCs) from various *Taxus* species, and after extensive research efforts involving media optimization, metabolic engineering and process engineering in both academic and industrial laboratories, Phyton Biotech, a DFB Pharmaceuticals company, received US Food and Drug Administration approval to produce Taxol for Bristol-Myers Squibb by plant cell culture⁴.

Nonetheless, this method of producing natural products has several limitations, including the slow growth rates of dedifferentiated

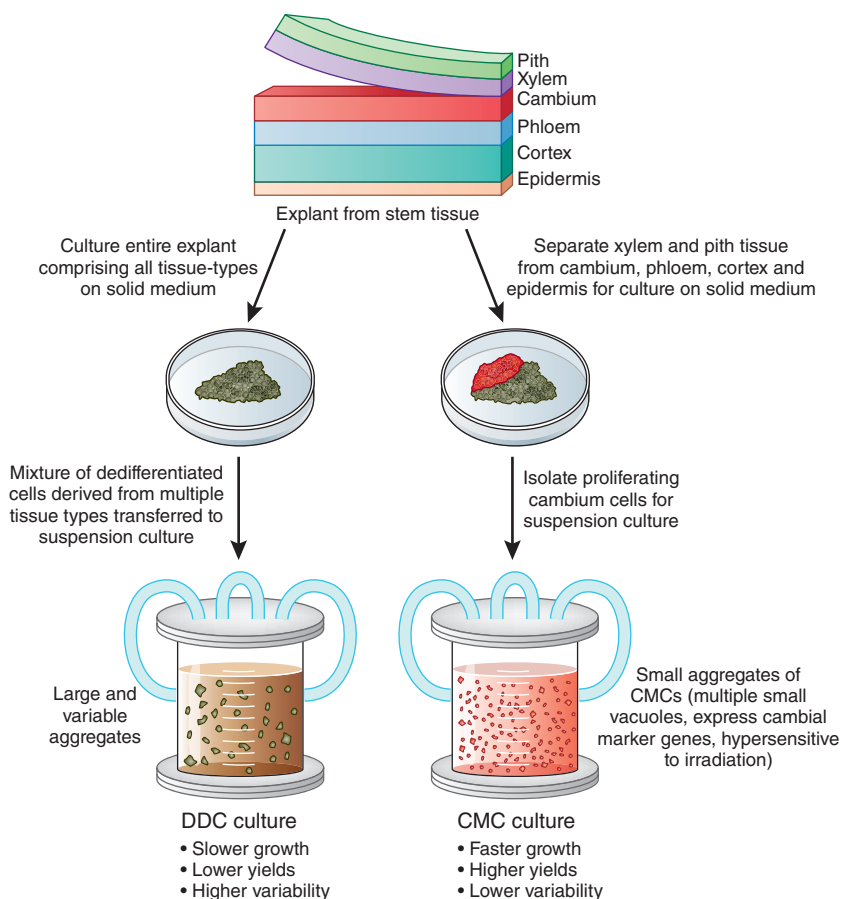


Figure 1 Suspension cultures of CMCs provide an attractive alternative to cultures of DDCs for the production of plant natural products.

plant cells, their aggregation (which complicates scale-up to bioreactors), low yields of secondary metabolites, and above all, variability in these properties⁵. Another alternative is microbial hosts engineered to express plant metabolic pathways. For example, *Escherichia coli* was recently engineered to express the first 2 out of the putative 19 steps of the dedicated paclitaxel biosynthetic pathway⁶. However, many complex plant secondary biosynthetic pathways are not fully defined, and it is unlikely

that this strategy will ensure a reliable and cost-effective supply of natural products such as paclitaxel in the near future. Currently, the best options involve a combination of approaches. Notable examples include production of the anti-malarial agent artemisinin by expression of a precursor in a microbial host followed by synthetic chemistry⁷, and a partial solution for paclitaxel that involves sustainable harvest of *Taxus* needles, extraction of a precursor and synthetic chemistry. In many situations,

Susan Roberts and Martin Kolewe are in the Chemical Engineering Department at the University of Massachusetts, Amherst, Massachusetts, USA.
e-mail: sroberts@ecs.umass.edu

Katie Vicari

however, such as natural products derived from slow-growing plants and having complex, undefined biosynthetic pathways, plant cell culture will remain the most viable option.

The work of Lee *et al.*³ marks an important departure from traditional plant cell culture. Instead of culturing heterogeneous mixtures of dedifferentiated cells, they isolated cells derived from vascular cambium and propagated them in solution. When exposed to the appropriate ratio of growth regulators, explants from most plant organs can be induced to dedifferentiate to form so-called callus cultures, which can be transferred to liquid media and disaggregated into single cells (Fig. 1). The resulting suspension cultures are amenable to bioprocessing technologies used for large-scale mammalian and microbial cultures. However, the starting cell population from the plant organ is a mixture of specialized cell types, which vary in cell cycle participation⁸ and other properties, and the cell types that remain after extended culture in the dedifferentiated state are also probably heterogeneous, contributing to the instability of many culture properties.

In contrast, Lee *et al.*³ bypass the dedifferentiation step by isolating undifferentiated CMCs from a variety of species, including *T. cuspidata*, *Panax ginseng* and *Ginkgo biloba*. The procedure (Fig. 1) is simple and rapid, with CMCs isolated from explants within a month. The authors focus on characterizing the potential of a cell line derived from *T. cuspidata* cambium cells for producing paclitaxel and related taxoids. The *T. cuspidata* CMC cultures are distinct from DDC cultures derived from needles and embryos on the basis of morphological characteristics, increased ability to differentiate, hypersensitivity to radiation and comparison of molecular signatures. A Gene Ontology analysis of the >500 differentially regulated genes confirms the upregulation of several genes known to be overexpressed in cambial cells. The cultured CMCs can thus be considered innately undifferentiated and inherently distinct from the dedifferentiated cells typically used for plant cell tissue culture.

The performance of CMCs as an applied bioprocessing technology is shown to be far superior to that of DDC cultures established at the same time. Growth rates and paclitaxel production are improved at both the laboratory scale (125 ml) and pilot scale (up to 20 liters), aggregate size is reduced, and variability in growth measured over repeated subculture cycles is markedly decreased. The authors also demonstrate the feasibility of using both stirred-tank and air-lift bioreactor designs for promoting growth and paclitaxel synthesis with CMCs, suggesting flexibility when designing a large-scale process. Although the

paclitaxel yields from CMCs are similar to some values reported elsewhere for dedifferentiated cells, they are lower than the maximum values reported⁹. However, Lee *et al.*³ did not attempt extensive process engineering optimization, and higher yields seem possible. The utility of this new approach is also demonstrated for the production of ginsenosides from *P. ginseng* cultures.

Some secondary metabolites accumulate to much higher extents in differentiated organ cultures (e.g., shoot or root). The most notable examples are the vinca alkaloids, produced by means of *Catharanthus roseus* hairy root cultures¹⁰. The benefits of using CMC cultures for the synthesis of such products will have to be evaluated.

A reliable, cost-effective supply of natural products for use as pharmaceuticals, fragrances, dyes and insecticides remains a major challenge for many systems. Plant cell tissue culture has been limited by inconsistent performance and the economic constraints associated with slow growth and low product yield. Compared with mammalian cell culture, plant cell culture has required batch times of months rather than weeks and has reached product titers of mg/l rather than g/l. Additional problems include cell aggregation, susceptibility to shearing,

variability in growth and profusion of necrotic cells. CMCs appear to enhance cell culture performance in all of these areas, and notably do not require selection of specific cells and aggregates for consistent growth over repeated subcultures, thereby minimizing maintenance requirements. These improvements yield cell cultures that are substantially closer to mammalian cell cultures with regard to large-scale process considerations. CMC-based strategies should therefore facilitate the development of economically viable plant cell tissue culture processes for many natural products.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Mann, J. *Nat. Rev. Cancer* **2**, 143–148 (2002).
2. Koehn, F.E. & Carter, G.T. *Nat. Rev. Drug Discov.* **4**, 206–220 (2005).
3. Lee, E.-K. *et al. Nat. Biotechnol.* **28**, 1213–1217 (2010).
4. Roberts, S.C. *Nat. Chem. Biol.* **3**, 387–395 (2007).
5. Kolewe, M.E., Gaurav, V. & Roberts, S.C. *Mol. Pharm.* **5**, 243–256 (2008).
6. Ajikumar, P.K. *et al. Science* **330**, 70–74 (2010).
7. Keasling, J.D. *ACS Chem. Biol.* **3**, 64–76 (2008).
8. Naill, M.C. & Roberts, S.C. *Biotechnol. Bioeng.* **90**, 491–500 (2005).
9. Bringi, V., Kadkade, P.G., Prince, C.L. & Roach, B.L. US patent 7264951 (2007).
10. Leonard, E., Runguphan, W., O'Connor, S. & Prather, K.J. *Nat. Chem. Biol.* **5**, 292–300 (2009).

Making antibodies from scratch

J Christopher Love

Synthesis and screening of a small library of antibody fragments yields promising hits.

Screening large combinatorial libraries is standard procedure in small-molecule drug discovery but is not generally used to discover therapeutic monoclonal antibodies. The main obstacle is the prohibitive cost of library synthesis, as it has been believed that upward of 100 million unique antibodies would be needed. In this issue, Mao *et al.*¹ overturn this conventional wisdom by showing that a library of only ~10,000 antibody fragments, designed *in silico* and synthesized *de novo*, can yield new leads for several targets. Their results suggest that rational approaches for determining the relationships between the sequences and activities of a library of

molecules can aid the search for new therapeutic antibodies.

Monoclonal antibodies are now widely used to treat a range of diseases, particularly cancers and autoimmune disorders. More than 20 are approved by the US Food and Drug Administration, generating over \$15 billion in revenue annually, and hundreds are in clinical trials². The insatiable demand for new therapeutic antibodies has spurred efforts to improve the efficiency of the processes used to discover them. The most common method used in industry relies on creating libraries of hybridomas from immunized (transgenic) mice, screening them by limiting serial dilution, and applying molecular engineering to humanize the leads or to improve their affinities and specificities. Advances in automated screening and liquid handling have improved this technique, but it remains arduous and costly.

J. Christopher Love is in the Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. e-mail: clove@mit.edu

however, such as natural products derived from slow-growing plants and having complex, undefined biosynthetic pathways, plant cell culture will remain the most viable option.

The work of Lee *et al.*³ marks an important departure from traditional plant cell culture. Instead of culturing heterogeneous mixtures of dedifferentiated cells, they isolated cells derived from vascular cambium and propagated them in solution. When exposed to the appropriate ratio of growth regulators, explants from most plant organs can be induced to dedifferentiate to form so-called callus cultures, which can be transferred to liquid media and disaggregated into single cells (Fig. 1). The resulting suspension cultures are amenable to bioprocessing technologies used for large-scale mammalian and microbial cultures. However, the starting cell population from the plant organ is a mixture of specialized cell types, which vary in cell cycle participation⁸ and other properties, and the cell types that remain after extended culture in the dedifferentiated state are also probably heterogeneous, contributing to the instability of many culture properties.

In contrast, Lee *et al.*³ bypass the dedifferentiation step by isolating undifferentiated CMCs from a variety of species, including *T. cuspidata*, *Panax ginseng* and *Ginkgo biloba*. The procedure (Fig. 1) is simple and rapid, with CMCs isolated from explants within a month. The authors focus on characterizing the potential of a cell line derived from *T. cuspidata* cambium cells for producing paclitaxel and related taxoids. The *T. cuspidata* CMC cultures are distinct from DDC cultures derived from needles and embryos on the basis of morphological characteristics, increased ability to differentiate, hypersensitivity to radiation and comparison of molecular signatures. A Gene Ontology analysis of the >500 differentially regulated genes confirms the upregulation of several genes known to be overexpressed in cambial cells. The cultured CMCs can thus be considered innately undifferentiated and inherently distinct from the dedifferentiated cells typically used for plant cell tissue culture.

The performance of CMCs as an applied bioprocessing technology is shown to be far superior to that of DDC cultures established at the same time. Growth rates and paclitaxel production are improved at both the laboratory scale (125 ml) and pilot scale (up to 20 liters), aggregate size is reduced, and variability in growth measured over repeated subculture cycles is markedly decreased. The authors also demonstrate the feasibility of using both stirred-tank and air-lift bioreactor designs for promoting growth and paclitaxel synthesis with CMCs, suggesting flexibility when designing a large-scale process. Although the

paclitaxel yields from CMCs are similar to some values reported elsewhere for dedifferentiated cells, they are lower than the maximum values reported⁹. However, Lee *et al.*³ did not attempt extensive process engineering optimization, and higher yields seem possible. The utility of this new approach is also demonstrated for the production of ginsenosides from *P. ginseng* cultures.

Some secondary metabolites accumulate to much higher extents in differentiated organ cultures (e.g., shoot or root). The most notable examples are the vinca alkaloids, produced by means of *Catharanthus roseus* hairy root cultures¹⁰. The benefits of using CMC cultures for the synthesis of such products will have to be evaluated.

A reliable, cost-effective supply of natural products for use as pharmaceuticals, fragrances, dyes and insecticides remains a major challenge for many systems. Plant cell tissue culture has been limited by inconsistent performance and the economic constraints associated with slow growth and low product yield. Compared with mammalian cell culture, plant cell culture has required batch times of months rather than weeks and has reached product titers of mg/l rather than g/l. Additional problems include cell aggregation, susceptibility to shearing,

variability in growth and profusion of necrotic cells. CMCs appear to enhance cell culture performance in all of these areas, and notably do not require selection of specific cells and aggregates for consistent growth over repeated subcultures, thereby minimizing maintenance requirements. These improvements yield cell cultures that are substantially closer to mammalian cell cultures with regard to large-scale process considerations. CMC-based strategies should therefore facilitate the development of economically viable plant cell tissue culture processes for many natural products.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Mann, J. *Nat. Rev. Cancer* **2**, 143–148 (2002).
2. Koehn, F.E. & Carter, G.T. *Nat. Rev. Drug Discov.* **4**, 206–220 (2005).
3. Lee, E.-K. *et al. Nat. Biotechnol.* **28**, 1213–1217 (2010).
4. Roberts, S.C. *Nat. Chem. Biol.* **3**, 387–395 (2007).
5. Kolewe, M.E., Gaurav, V. & Roberts, S.C. *Mol. Pharm.* **5**, 243–256 (2008).
6. Ajikumar, P.K. *et al. Science* **330**, 70–74 (2010).
7. Keasling, J.D. *ACS Chem. Biol.* **3**, 64–76 (2008).
8. Naill, M.C. & Roberts, S.C. *Biotechnol. Bioeng.* **90**, 491–500 (2005).
9. Bringi, V., Kadkade, P.G., Prince, C.L. & Roach, B.L. US patent 7264951 (2007).
10. Leonard, E., Runguphan, W., O'Connor, S. & Prather, K.J. *Nat. Chem. Biol.* **5**, 292–300 (2009).

Making antibodies from scratch

J Christopher Love

Synthesis and screening of a small library of antibody fragments yields promising hits.

Screening large combinatorial libraries is standard procedure in small-molecule drug discovery but is not generally used to discover therapeutic monoclonal antibodies. The main obstacle is the prohibitive cost of library synthesis, as it has been believed that upward of 100 million unique antibodies would be needed. In this issue, Mao *et al.*¹ overturn this conventional wisdom by showing that a library of only ~10,000 antibody fragments, designed *in silico* and synthesized *de novo*, can yield new leads for several targets. Their results suggest that rational approaches for determining the relationships between the sequences and activities of a library of

molecules can aid the search for new therapeutic antibodies.

Monoclonal antibodies are now widely used to treat a range of diseases, particularly cancers and autoimmune disorders. More than 20 are approved by the US Food and Drug Administration, generating over \$15 billion in revenue annually, and hundreds are in clinical trials². The insatiable demand for new therapeutic antibodies has spurred efforts to improve the efficiency of the processes used to discover them. The most common method used in industry relies on creating libraries of hybridomas from immunized (transgenic) mice, screening them by limiting serial dilution, and applying molecular engineering to humanize the leads or to improve their affinities and specificities. Advances in automated screening and liquid handling have improved this technique, but it remains arduous and costly.

J. Christopher Love is in the Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. e-mail: clove@mit.edu

The ability to clone antibody genes by PCR has led to a second approach in which *in vitro* libraries of recombinant antibodies are displayed on the surfaces of organisms such as phage, yeast or bacteria, and the highest-affinity binders are isolated by selection³. Such libraries of antibodies (or fragments of their binding regions) are usually designed to maximize their size and thus their diversity. This strategy has been motivated by (i) the theoretical diversity of natural antibody repertoires, estimated at $>10^{10}$ based on the number of unique combinations of gene segments possible and subsequent additions and deletions⁴, (ii) the breadth of diversity feasible for libraries of genes in suitable hosts (10^7 – 10^{12}) (ref. 3) and (iii) the general unimportance of library size in selection-based strategies. Although displayed recombinant libraries are now widely used for discovery, selections typically rely on rounds of enrichment based only on affinities. It is not clear, however, that it is always desirable to identify antibodies with the highest affinities for a target^{5,6}.

Natural antibody repertoires may in fact be smaller than the predicted diversity of 10^{10} . Mice whose antibody repertoires are constrained by a single gene encoding the variable region of the heavy chain are still able to produce specific antibodies to most antigens⁷. Moreover, high-throughput sequencing of the variable genes expressed by zebrafish B cells has shown that only a relatively small percentage of the possible combinations of genes account for a large fraction of the repertoire in an individual⁸. Similar analysis of the diversity in circulating human B cells has indicated that the numbers of unique clones may be as low as 10^6 (ref. 9). These findings, among others, suggest that relatively small libraries could yield interesting monoclonal antibodies.

Mao *et al.*¹ have now shown this idea to be true experimentally. They computationally designed a combinatorial library of heavy and light chain antibody germline genes that represented a diverse structural repertoire and then *de novo* synthesized DNA encoding the genes (Fig. 1). Antibody fragments were constructed by transforming *Escherichia coli* with heavy and light chain genes, and then purifying the expressed protein in individual wells of 96-well microtiter plates. This process yielded a defined, spatially addressable library of ~10,000 members that could be screened using conventional high-throughput screening tools—that is, the sequence of every candidate molecule is known beforehand, unlike in hybridoma- or display-based methods. Remarkably, the small library yielded 85 hits for 7 out of 9 antigens screened in parallel, with affinities ranging from ~0.5 μM to 100 μM .

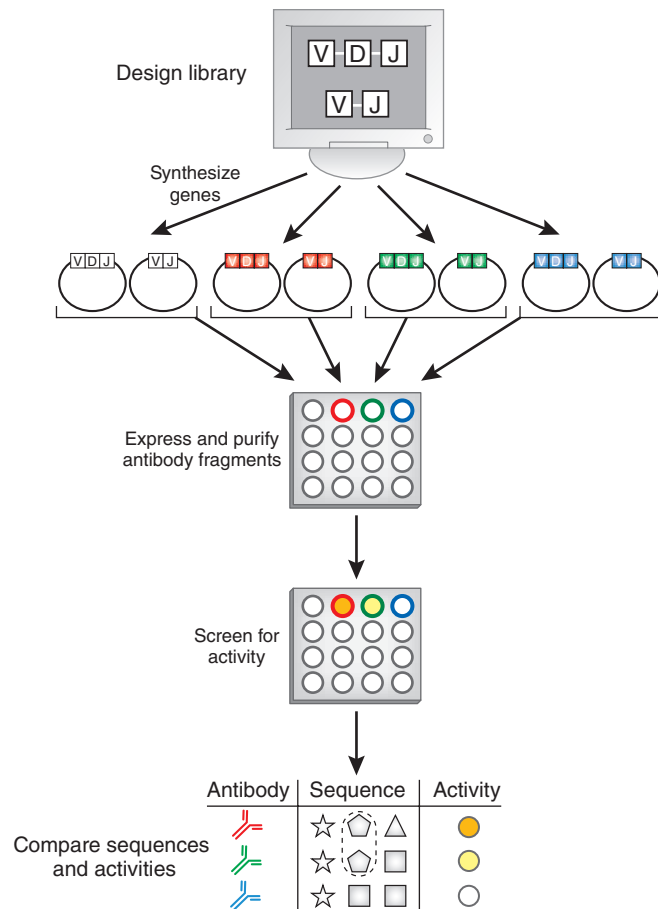


Figure 1 Strategy for the design, synthesis and screening of a spatially addressable library of antibody fragments. Members of the library are designed *in silico* by defining specific combinations of germline gene segments for the variable regions of the heavy and light chains of an antibody. Each pair of genes encoding a unique combination of heavy and light chains is synthesized *de novo* and used to transform *E. coli* in a known location in a microtiter plate. After parallel expression and purification, the antibody fragments are screened for specificity to nine unique antigens. Information about binding for each member in the library can then be related to their sequences to highlight key positions in the sequences that confer specificity. These sites become targets for subsequent modifications to alter the antibody's specificity, affinity or biological activity.

Because the sequence of each member in the library was known *a priori*, these hits could be compared to other members in the library having similar sequences but different affinities or specificities. This correlation allowed direct assessments of the 'sequence-activity' relationships in a manner analogous to relating structures and activities for small-molecule drugs. Essential residues highlighted by this analysis were then targeted for subsequent maturation to improve affinities and tune biological activity. Mao *et al.*¹ describe one matured, full-length antibody that partially antagonized Notch-1 signaling by blocking delta-like ligand 4 (DLL-4) and another antibody that exhibited more potent inhibition. These results show that a single screen could yield several lead structures with distinct biological activities.

Some targets will present challenges for the approach of Mao *et al.*¹ Antibodies expressed

from germline genes have low affinities naturally, and the immune system relies on avidity to stabilize weak interactions with antigens. Naive B cells express many copies of their unique B-cell receptors on their surface, and natural decavalent IgM circulates to bind new antigens into immune complexes. In some cases, the affinities of monovalent antibody fragments may simply be too low to score. Generating libraries of multivalent species may help address this issue. Nonetheless, certain targets or functional activities may still pose difficulties. For example, all broadly neutralizing antibodies against HIV-1 identified to date have unusual structural elements (e.g., domain swaps and large loops in the complementarity-determining regions) arising from significant maturation *in vivo*¹⁰. It is not obvious that simple 'sequence-activity' analysis for germline genes would reveal the

Katie Vicari

best candidates for such antibodies or the subsequent routes to maturation.

Further technical developments of the method described by Mao *et al.*¹ should accommodate larger libraries than those used here while maintaining specific knowledge about each member. Declining costs of gene synthesis and advancing systems for miniaturized expression, purification and screening of multiple recombinant proteins in parallel should also help. Such advances would begin to bring the vast experience available in high-throughput screening of combinatorial libraries of small molecules to bear on the problem of identifying new therapeutic monoclonal antibodies.

COMPETING FINANCIAL INTERESTS

The author declares no competing financial interests.

1. Mao, H. *et al.* *Nat. Biotechnol.* **28**, 1195–1202 (2010).
2. Nelson, A.L., Dhimolea, E. & Reichert, J.M. *Nat. Rev. Drug Discov.* **9**, 767–774 (2010).
3. Hoogenboom, H.R. *Nat. Biotechnol.* **23**, 1105–1116 (2005).
4. Kindt, T.J., Osborne, B.A. & Goldsby, R.A. *Kuby Immunology*, edn. 6 (W.H. Freeman, 2006).
5. Suntharalingam, G. *et al.* *N. Engl. J. Med.* **355**, 1018–1028 (2006).
6. Weiner, L.M. & Carter, P. *Nat. Biotechnol.* **23**, 556–557 (2005).
7. Xu, J.L. & Davis, M.M. *Immunity* **13**, 37–45 (2000).
8. Weinstein, J.A., Jiang, N., White, R.A., Fisher, D.S. & Quake, S.R. *Science* **324**, 807–810 (2009).
9. Boyd, S.D. *et al.* *Sci. Transl. Med.* **1**, 12ra23 (2009).
10. Kwong, P.D. & Wilson, I.A. *Nat. Immunol.* **10**, 573–578 (2009).

regions (UTRs). The start and stop codons, which define the protein open reading frame, tend to occur in relatively unstructured regions, a feature that may facilitate access by the translation machinery. The relative lack of secondary structure in the 3' UTR may ensure that factors that regulate mRNA translation, stability and localization have access to their binding sites.

The finding that open reading frames are more structured than the flanking regions is consistent with a study from our laboratory³ showing that the structured regions of the HIV-1 genome occur preferentially downstream of sequences encoding individual protein domains, perhaps facilitating protein folding during translation. Also in accordance with earlier results⁴, Kertesz *et al.*² found that highly translated messages appear to have less defined structure near the translational start site than those with lower translation levels, suggesting that mRNA structure modulates ribosome activity.

Characterizing the extent of double- and single-stranded structure in the yeast transcriptome represents an important first step toward the goal of determining the complete three-dimensional structures of these RNAs. However, generating accurate nucleotide-resolution structures from RNase cleavage data is very challenging⁵. Moreover, *in vitro* folding, as used in PARS, is not likely to fully recapitulate folding in the cellular environment. The ability of any approach to guide RNA structure prediction is best assessed by comparing predicted models to RNA structures known to be in a functional conformation. Although very few yeast RNAs have been studied over significant lengths, the secondary structure of the 18S ribosomal RNA from the small (40S) ribosomal subunit is well characterized⁶. We found that the PARS-assisted, secondary-structure prediction for the 18S RNA² contains only half of the accepted base pairs for this RNA and thus does not resemble the biologically functional state in many regions. In addition, the PARS-assisted secondary structure for 9 of 14 tRNAs with read depths greater than 1 were predicted incorrectly. Thus, at this juncture, PARS results do not reproduce these physiologically relevant RNA structures and should be interpreted cautiously.

In contrast, previous, lower-throughput approaches have yielded very high-resolution structural information under defined, controlled conditions in which the RNAs retain most of their native structure. In the mid-1980s, Harry Noller's laboratory demonstrated that primer extension could identify RNA nucleotides modified by small, structure-selective, organic chemicals. Using these reagents,

Toward global RNA structure analysis

David M Mauger & Kevin M Weeks

Deep sequencing provides a first view of the RNA structures in a eukaryotic transcriptome.

The structures of RNA molecules, even in complex environments, have been interrogated for many years by studying how individual nucleotides react with enzymatic or chemical probes¹. However, these methods have largely been limited to studying single RNAs or small fragments of large RNAs. Writing in *Nature*, Kertesz *et al.*² have succeeded in analyzing much of the yeast transcriptome by melding a traditional biochemical method for probing RNA structure with the power of highly parallel DNA sequencing in an approach called parallel analysis of RNA structure (PARS). Their results provide new insight into the role of mRNA structure in gene expression, confirming, in particular, that RNA structure regulates protein synthesis, most probably by controlling the accessibility of mRNAs to the translational machinery.

Kertesz *et al.*² extracted total RNA from *Saccharomyces cerevisiae*, enriched it for mRNA, refolded the RNA, and treated it separately with S1 nuclease and RNase V1, nucleases specific for single- and double-stranded RNA, respectively (Fig. 1). The resulting

structure-selective cleavage fragments were converted to double-stranded DNAs and sequenced using the SOLiD system, generating tens of millions of reads. Sequences within less abundant transcripts and those that were more difficult to convert from RNA to double-stranded DNA were largely undetected. The authors then determined individual RNase cleavage sites from the sequencing reads and compared the digestion frequencies of the two RNases at each nucleotide. This comparison, termed a PARS score, provides a measure of the single-stranded or double-stranded character of each nucleotide in each RNA with sufficient read coverage. By characterizing thousands of yeast transcripts, the authors generated structural data for approximately half of the *S. cerevisiae* transcriptome.

These data are especially valuable when analyzed at a global level (Fig. 1). In aggregate, they lend strong support to the general hypothesis that the information encoded in large mRNAs is regulated and organized by RNA structure. In several cases, the authors were able to confirm and extend models for the roles of RNA structure previously developed through detailed analyses of individual RNAs. The data revealed that the prototypical yeast mRNA consists of regions with distinct characteristics: the open reading frame is more structured than either the 5' or 3' untranslated

David M. Mauger and Kevin M. Weeks are in the Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina, USA.

e-mail: weeks@unc.edu

best candidates for such antibodies or the subsequent routes to maturation.

Further technical developments of the method described by Mao *et al.*¹ should accommodate larger libraries than those used here while maintaining specific knowledge about each member. Declining costs of gene synthesis and advancing systems for miniaturized expression, purification and screening of multiple recombinant proteins in parallel should also help. Such advances would begin to bring the vast experience available in high-throughput screening of combinatorial libraries of small molecules to bear on the problem of identifying new therapeutic monoclonal antibodies.

COMPETING FINANCIAL INTERESTS

The author declares no competing financial interests.

1. Mao, H. *et al.* *Nat. Biotechnol.* **28**, 1195–1202 (2010).
2. Nelson, A.L., Dhimolea, E. & Reichert, J.M. *Nat. Rev. Drug Discov.* **9**, 767–774 (2010).
3. Hoogenboom, H.R. *Nat. Biotechnol.* **23**, 1105–1116 (2005).
4. Kindt, T.J., Osborne, B.A. & Goldsby, R.A. *Kuby Immunology*, edn. 6 (W.H. Freeman, 2006).
5. Suntharalingam, G. *et al.* *N. Engl. J. Med.* **355**, 1018–1028 (2006).
6. Weiner, L.M. & Carter, P. *Nat. Biotechnol.* **23**, 556–557 (2005).
7. Xu, J.L. & Davis, M.M. *Immunity* **13**, 37–45 (2000).
8. Weinstein, J.A., Jiang, N., White, R.A., Fisher, D.S. & Quake, S.R. *Science* **324**, 807–810 (2009).
9. Boyd, S.D. *et al.* *Sci. Transl. Med.* **1**, 12ra23 (2009).
10. Kwong, P.D. & Wilson, I.A. *Nat. Immunol.* **10**, 573–578 (2009).

regions (UTRs). The start and stop codons, which define the protein open reading frame, tend to occur in relatively unstructured regions, a feature that may facilitate access by the translation machinery. The relative lack of secondary structure in the 3' UTR may ensure that factors that regulate mRNA translation, stability and localization have access to their binding sites.

The finding that open reading frames are more structured than the flanking regions is consistent with a study from our laboratory³ showing that the structured regions of the HIV-1 genome occur preferentially downstream of sequences encoding individual protein domains, perhaps facilitating protein folding during translation. Also in accordance with earlier results⁴, Kertesz *et al.*² found that highly translated messages appear to have less defined structure near the translational start site than those with lower translation levels, suggesting that mRNA structure modulates ribosome activity.

Characterizing the extent of double- and single-stranded structure in the yeast transcriptome represents an important first step toward the goal of determining the complete three-dimensional structures of these RNAs. However, generating accurate nucleotide-resolution structures from RNase cleavage data is very challenging⁵. Moreover, *in vitro* folding, as used in PARS, is not likely to fully recapitulate folding in the cellular environment. The ability of any approach to guide RNA structure prediction is best assessed by comparing predicted models to RNA structures known to be in a functional conformation. Although very few yeast RNAs have been studied over significant lengths, the secondary structure of the 18S ribosomal RNA from the small (40S) ribosomal subunit is well characterized⁶. We found that the PARS-assisted, secondary-structure prediction for the 18S RNA² contains only half of the accepted base pairs for this RNA and thus does not resemble the biologically functional state in many regions. In addition, the PARS-assisted secondary structure for 9 of 14 tRNAs with read depths greater than 1 were predicted incorrectly. Thus, at this juncture, PARS results do not reproduce these physiologically relevant RNA structures and should be interpreted cautiously.

In contrast, previous, lower-throughput approaches have yielded very high-resolution structural information under defined, controlled conditions in which the RNAs retain most of their native structure. In the mid-1980s, Harry Noller's laboratory demonstrated that primer extension could identify RNA nucleotides modified by small, structure-selective, organic chemicals. Using these reagents,

Toward global RNA structure analysis

David M Mauger & Kevin M Weeks

Deep sequencing provides a first view of the RNA structures in a eukaryotic transcriptome.

The structures of RNA molecules, even in complex environments, have been interrogated for many years by studying how individual nucleotides react with enzymatic or chemical probes¹. However, these methods have largely been limited to studying single RNAs or small fragments of large RNAs. Writing in *Nature*, Kertesz *et al.*² have succeeded in analyzing much of the yeast transcriptome by melding a traditional biochemical method for probing RNA structure with the power of highly parallel DNA sequencing in an approach called parallel analysis of RNA structure (PARS). Their results provide new insight into the role of mRNA structure in gene expression, confirming, in particular, that RNA structure regulates protein synthesis, most probably by controlling the accessibility of mRNAs to the translational machinery.

Kertesz *et al.*² extracted total RNA from *Saccharomyces cerevisiae*, enriched it for mRNA, refolded the RNA, and treated it separately with S1 nuclease and RNase V1, nucleases specific for single- and double-stranded RNA, respectively (Fig. 1). The resulting

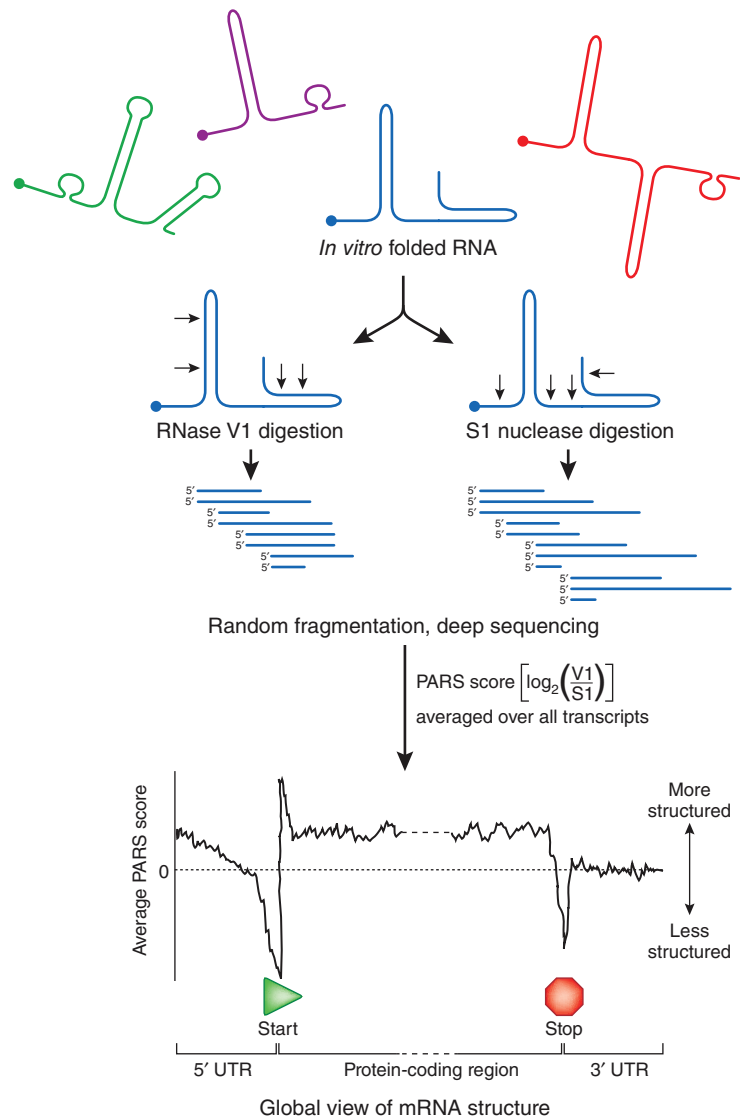
structure-selective cleavage fragments were converted to double-stranded DNAs and sequenced using the SOLiD system, generating tens of millions of reads. Sequences within less abundant transcripts and those that were more difficult to convert from RNA to double-stranded DNA were largely undetected. The authors then determined individual RNase cleavage sites from the sequencing reads and compared the digestion frequencies of the two RNases at each nucleotide. This comparison, termed a PARS score, provides a measure of the single-stranded or double-stranded character of each nucleotide in each RNA with sufficient read coverage. By characterizing thousands of yeast transcripts, the authors generated structural data for approximately half of the *S. cerevisiae* transcriptome.

These data are especially valuable when analyzed at a global level (Fig. 1). In aggregate, they lend strong support to the general hypothesis that the information encoded in large mRNAs is regulated and organized by RNA structure. In several cases, the authors were able to confirm and extend models for the roles of RNA structure previously developed through detailed analyses of individual RNAs. The data revealed that the prototypical yeast mRNA consists of regions with distinct characteristics: the open reading frame is more structured than either the 5' or 3' untranslated

David M. Mauger and Kevin M. Weeks are in the Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina, USA.

e-mail: weeks@unc.edu

Figure 1 Overview of the PARS approach for measuring RNA structure on a transcriptome scale. RNAs are refolded *in vitro*, cleaved with structure-selective ribonucleases, fragmented and subjected to highly parallel sequencing. The PARS score gives a measure of double versus single strandedness. One use of the PARS scores is to average these values over many RNAs to yield a global view of RNA structure. PARS suggests that most *S. cerevisiae* transcripts have less secondary structure in untranslated regions (UTRs) than in coding regions, and that mRNAs with structured translation start sites are translated with lower efficiency than those with unstructured translation start sites.



Noller and colleagues examined the structure of the *Escherichia coli* 16S ribosomal RNA and detected fine-scale perturbations resulting from the binding of proteins and drugs^{7,8}. We developed a descendant of this technology, called selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE), which makes use of broadly reactive, yet structure-selective, small-molecule reagents. Using SHAPE, it is often possible to predict RNA structures with high accuracy⁹.

As next-generation technologies for probing RNA structure are further refined, it will be important to combine such innovations in high-throughput data acquisition with bioinformatic approaches for accurately interpreting complex data sets. Eventually, it may be possible to carry out global analyses of how RNA structures change in response to environmental stimuli and developmental cues. The ability to accurately determine the structures of and interactions between components of entire transcriptomes will contribute substantially to our understanding of the biological roles of RNA, including in RNA processing, translation and non-coding RNA function. This understanding should facilitate the use of molecular biology tools to manipulate the cellular state and develop RNA-directed therapeutics.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Zaug, A.J. & Cech, T.R. *RNA* **1**, 363–374 (1995).
2. Kertesz, M. *et al.* *Nature* **467**, 103–107 (2010).
3. Watts, J.M. *et al.* *Nature* **460**, 711–716 (2009).
4. Kozak, M. *Gene* **361**, 13–37 (2005).

5. Mathews, D.H., Sabina, J., Zuker, M. & Turner, D.H. *J. Mol. Biol.* **288**, 911–940 (1999).
6. Gutell, R.R. *Nucleic Acids Res.* **21**, 3051–3054 (1993).
7. Moazed, D., Stern, S. & Noller, H.F. *J. Mol. Biol.* **187**, 399–416 (1986).
8. Moazed, D. & Noller, H.F. *Nature* **327**, 389–394 (1987).
9. Deigan, K.E., Li, T.W., Mathews, D.H. & Weeks, K.M. *Proc. Natl. Acad. Sci. USA* **106**, 97–102 (2009).

In silico epitope scaffolds

A protein's ability to induce antibodies resides in the structure of individual epitopes, with some more effective at eliciting neutralizing antibodies than others. Ofek *et al.* describe a method for creating and optimizing epitopes on scaffolds, predicted *in silico*, that can be used to generate high-affinity antibodies *in vivo*. Using as a model an epitope from the membrane-proximal external region of HIV-1 gp41 protein, they first searched a protein database for scaffolds that structurally resemble the epitope bound to an antibody (2F5) directed against it, and introduced mutations to improve scaffold-epitope binding. Having designed five scaffolds *in silico*, they analyzed the biophysical interaction of each with 2F5 antibody. They found that they all bound 2F5 with roughly the same affinity as free peptide, but that there were differences among the five scaffolds in how flexible they were, which interestingly translated into differences in their ability to elicit an immune response *in vivo*. This method for recreating and optimizing epitopes may provide a way to generate antibodies to recalcitrant, yet medically important, proteins. In a related study, some of the same researchers achieved similar success with the poorly immunogenic HIV epitope 4E10, another membrane-proximal epitope in gp41. This system may also provide useful insights into the adaptive immune response to particular structures. (*Proc. Natl. Acad. Sci. USA*, published online 27 September 2010; doi:10.1073/pnas.1004728107; *Structure* **18**, 1116–1126, 2010)



LD

Identifying transcription factor modulators

How does genetic variation affect gene expression? Approaches to date have used genetic linkage analysis to associate genetic markers with the expression of genes measured across a population of closely related individuals. In a twist on these approaches, Lee and Bussemaker first infer the 'activity' of transcription factors in each individual and then look for markers linked to transcription factor activity. The authors posit that this extra computational step increases the signal-to-noise ratio by combining into a single metric the expression of many genes predicted to be targets of a transcription factor. Transcription factor activity was inferred from gene expression profiles, the promoter sequence of each gene and known DNA-binding specificities of transcription factors. When applied to data from yeast, the approach outperformed previous methods, identifying over six times as many associations between transcription factors and loci that may modulate their activity. Although some of these links are consistent with existing data and could be supported experimentally, much work remains to understand the importance of each association. (*Mol. Syst. Biol.* **6**, 412, 2010)

CM

Deep sequencing and RNA fitness

The vastness of sequence space—a 20-mer RNA, for example, has $\sim 10^{12}$ possible sequences—makes it extremely challenging to visualize the relationships between all possible genotypes of a macromolecule and their cognate phenotypes. Pitt and Ferré-D'Amaré demonstrate the potential of next-generation sequencing methods to profile the activities associated with different RNA sequences—so-called fitness landscapes. Working with a 54-nucleotide RNA ligase ribozyme, they use deep sequencing to characterize both the sequences and activities (measured as abundances) of a population of variants of the catalytic RNA both before and after *in vitro* selection. By projecting the empirically determined fitness onto the ribozyme sequence, they are able to identify functionally important residues not previously known to be required for maximal activity. Besides its potential use in such applications as the optimization of aptamer sequences, the same principle could theoretically be used to elucidate the molecular basis of pathogen fitness in patient populations, which might in turn have implications for vaccine and drug development. (*Science* **330**, 376–379, 2010)

PH

MRI for microfluidics

The potential of microfluidics to increase the throughput and portability of analyses has been limited by the need for a tool that can probe both microscale chemistry and flow dynamics. Conventional magnetic resonance imaging (MRI) cannot sensitively record images of microfluidic flow with the required spatial and temporal resolution. By adapting remotely detected MRI—a variant of conventional MRI in which the signal-encoding phase of MRI is decoupled from the signal detection process—and combining it with JPEG-style compressive sampling, Bajaj *et al.* are able to dramatically increase the speed of acquisition of MRI images. The spatial resolution they achieve is sufficient to capture the results of up to several thousand parallel assays on a microfluidic device. Although the approach is only relevant to nuclear magnetic resonance (NMR)-active analytes or those that can be detected indirectly by contrast agents or other sensors, it may open new opportunities for a range of highly parallel analytical applications involving microfluidic chips. (*Science*, published online 7 October 2010; doi:10.1126/science.1192313)

PH

Lentivectors and clinical validation

Gene-therapy vectors based on lentiviruses may be safer than gammaretrovirus vectors, which have caused leukemia in several patients by integrating into the genome near proto-oncogenes. Last year, the first reported clinical trial of a lentiviral vector found no evidence of cancer or clonal cell amplification in two children with adrenoleukodystrophy. A second clinical trial based on a lentiviral vector has now released results on one patient. This study treated beta-thalassemia, a disease of red blood cells, with autologous hematopoietic stem cells genetically modified to express the missing beta-globin chain of hemoglobin. One year after the transplant, the patient's red blood cells had recovered sufficiently that he has not required blood transfusions for 28 months. But the therapy did elicit a clonal amplification: a cell carrying a vector insertion in *HMGA2*, a gene involved in transcriptional activation, expanded to $\sim 50\%$ of transduced cells, although staying below 6% of all blood cells. Why this occurred remains to be determined. Overexpression of full-length and truncated *HMGA2* has been linked to benign and malignant tumors, and, about three years after the transplant, it is unclear whether the clonal expansion seen here is innocuous or "a prelude to multistep leukaemogenesis?" (*Nature* **467**, 318–322, 2010)

KA

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

In silico research in the era of cloud computing

Joel T Dudley & Atul J Butte

Snapshots of computer systems that are stored and shared 'in the cloud' could make computational analyses more reproducible.

Scientific findings must be reproducible for them to be formally accepted by colleagues, practitioners, policy makers and the lay public. Although it was once thought that computers would improve reproducibility because they yield repeatable results given the same set of inputs, most software tools do not provide mechanisms to package a computational analysis such that it can be easily shared and reproduced. This inability to easily exchange the computational analyses behind published results hinders reproducibility across scientific disciplines¹.

Several innovative solutions to this problem have been proposed. Most efforts have focused on software tools that aim to standardize the creation, representation and sharing of computational 'workflows'² that tie several software tools together into a single analysis. These workflow tools provide a way to represent discrete computational tasks (e.g., processing an input data file) as computational modules that can be connected into workflows by linking the output of one module with the input of another (e.g., the output from the input-data-processing module connects to the input of a data normalization module)³. Several of these tools allow researchers to build complex computational workflows through drag-and-drop visual interfaces and to share standardized representations of workflows. Examples include

GenePattern⁴ for analyzing genomic data, the Trident Scientific Workflow Workbench for oceanography (<http://www.microsoft.com/mscorp/tc/trident.msp>), Taverna⁵ for generalized scientific computation and web-based resources such as myExperiment⁶ for sharing workflows. Recently, a software solution has been described⁷ that embeds access to computational systems directly into digital representations of scientific papers.

Existing software applications have not become established solutions to the problem of computational reproducibility. This is not because of any technical shortcomings of the software because, in fact, many programs are technically well designed. Rather, the failures are a consequence of human nature and the realities of data-driven science, including the following: first, efforts are not rewarded by the current academic research and funding environment^{8,9}; second, commercial software vendors tend to protect their markets through proprietary formats and interfaces¹⁰; third, investigators naturally tend to want to own and control their research tools; fourth, even the most generalized software will not be able to meet the needs of every researcher in a field; and finally, the need to derive and publish results as quickly as possible precludes the often slower standards-based development path⁹.

The consequence is that nonstandardized, research computational pipelines continue to be developed, especially as new types of molecular measurements generate quantities of data that most standardized software packages cannot handle¹¹. We can thus conclude that any effort to establish standard protocols that require investigators to adopt a particular software system, creating a real or perceived

threat of 'lock in' to software from a single laboratory or commercial vendor, will be met with disregard or even resistance and will likely fail to drastically change the current behavior of researchers. Given these realities, we propose capturing and exchanging computational pipelines using complete digital representations of the entire computing environment needed to execute the pipeline.

Whole system snapshot exchange

In this approach, which we call whole system snapshot exchange (WSSE), the computer system(s) used by researchers to produce experimental results are copied in their entirety, including the operating system, application software and databases, into a single digital image that can be exchanged with other researchers. Using WSSE, researchers would be able to obtain precise replicas of a computational system used to produce the published results and have the ability to restore this system to the precise state of the system when the experimental results were generated. Even subtle variances caused by discrepancies between specific versions of software or programming languages are avoided by the WSSE approach. In principle, the input data used to produce published results could be exchanged along with the pipeline.

Readers may be quick to realize that WSSE could involve the exchange of data files that might range in size from tens of gigabytes to several terabytes or more. Even with the fastest internet connections, it is not feasible to share such large files easily and efficiently. Thus, it is not our intention that WSSE operate desktop-to-desktop, but rather we propose that data and computational pipelines

Joel T. Dudley and Atul J. Butte are in the Division of Systems Medicine, Department of Pediatrics, Stanford University School of Medicine, Stanford, California, USA, and at the Lucile Packard Children's Hospital, Palo Alto, California, USA.
e-mail: abutte@stanford.edu

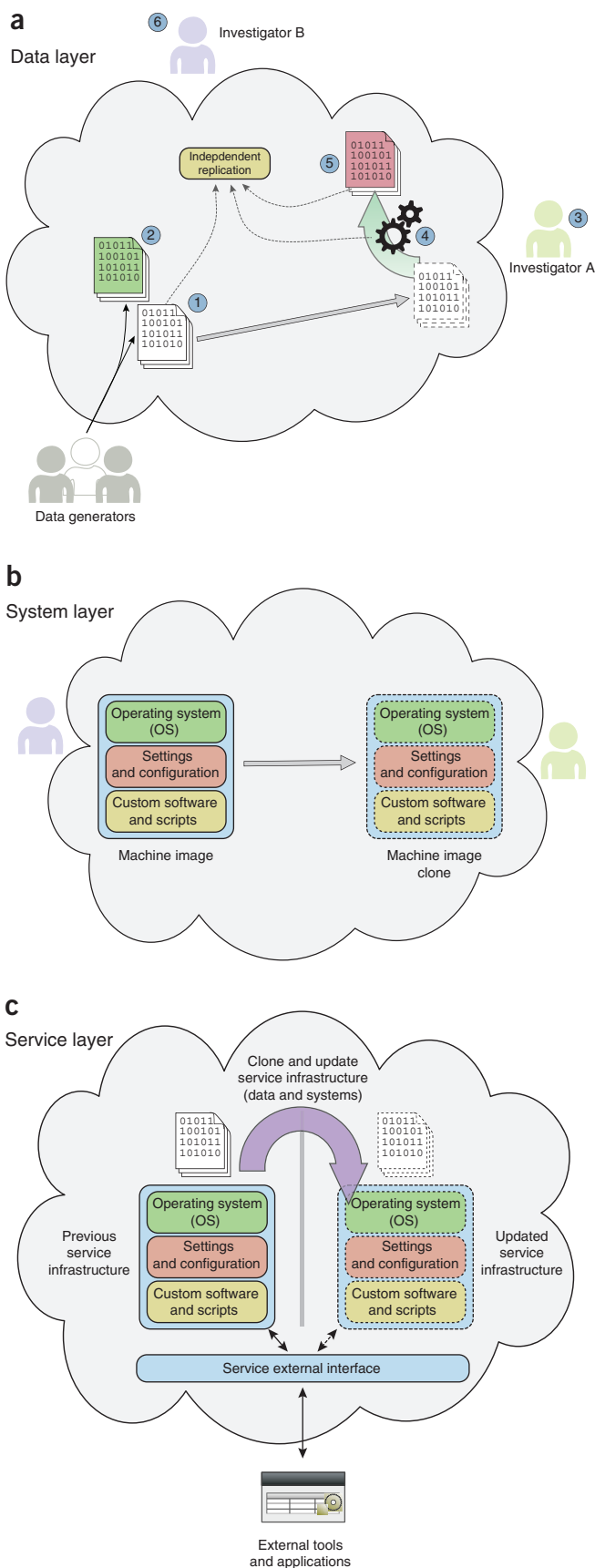


Figure 1 Layers of reproducible computing in the cloud. **(a)** Data layer. Generators of large scientific data sets can publish their data to the cloud (1) and substantial updates to these data sets can exist in parallel without loss or modification of the previous data set (2). Primary investigators can clone entire data sets within the cloud (3) and apply custom scripts or software computations (4) to derive published results (5). An independent investigator can obtain digital replicates of the original primary data set, software and published results within the cloud to replicate a published analysis and compare it with published results (6). **(b)** System layer. Investigators can set up and conduct scientific computations using cloud-based virtual machine images that contain all of the software, configurations and scripts necessary to execute the analysis. The customized machine image can be copied and shared with other investigators within the cloud for replicate analyses. **(c)** Service layer. Instead of replacing an old version of a scientific computing service with a new version of the software, a service that is virtualized in the cloud can be easily replicated and the old version made available alongside the new. Requests made by applications through the external service interface could incorporate a version parameter. This could enable published results that used older versions of the service to be evaluated for reproducibility.

will be exchanged exclusively using cloud computing, defined here as “computing in which dynamically scalable and virtualized resources are provided as a service over the Internet”¹².

In cloud computing, computation and data are ‘virtualized’, meaning that software and data are not tied to physical computing resources, such as a specific server with hard drives plugged into it. Instead, cloud-computing infrastructures are comprised of large and often geographically disparate clusters of computing hardware that are made to appear as a single, homogeneous computational environment¹³. The virtualization of data in the cloud makes it possible to move or copy ‘snapshots’ of large data sets from point to point within the cloud at high transfer rates, without the need to associate particular machines or storage drives at either the source or destination of the data transfer. Some cloud providers offer technologies that enable the creation of large databases that can persist as omnipresent resources that can be accessed by any computation in the cloud.

Concerns have been voiced¹⁴ that scientific computing in the cloud could make results less reproducible. One concern is that cloud computing will be a computing ‘black box’ that obfuscates details needed to accurately interpret the results of computational analyses. Although this is an important concern, we argue that cloud computing could actually

make scientific computing more transparent by making it practical to share entire computational infrastructures so that they can be scrutinized. Without cloud computing, in contrast, sharing research software and data typically requires that they be modified and repackaged for distribution, and whoever receives the software must possess the necessary computing infrastructure.

Cloud-based support for reproducibility

Cloud computing could support reproducibility in several ways, which correspond to different 'layers' for accessing computing resources (Fig. 1).

At the data layer, cloud computing enables data sets to be easily stored and shared virtually—that is, without necessarily copying it to another computer. Most importantly, when a data set has been used to derive published results, it can be copied and archived in the cloud. This would be facilitated if large public data repositories were regularly archived into the cloud. One cloud-computing vendor, Amazon Web Services, has already taken the initiative to archive many such public data sets. Their extensive online catalog (<http://aws.amazon.com/publicdatasets/>) contains large public data sets from various domains of science, including astronomy, biology, chemistry and climatology.

At the system layer, cloud computing allows snapshots of complete computer systems to be exchanged, as proposed in WSSE. This addresses observations that computer systems can substantially confound the reproducibility of analyses¹⁵.

A higher-level layer of scientific computation in the cloud is the service layer, in which computational services are exposed to external applications through some form of programming interface¹⁶. Examples include the Entrez Utilities from NCBI (<http://eutils.ncbi.nlm.nih.gov/>) that provide access to data and computational resources from NCBI's repertoire of bioinformatics applications and databases. From the standpoint of reproducibility, there are problems introduced by such a service-oriented approach to scientific computing: the underlying application supporting the computational service may be altered significantly without any apparent change to the public-facing service interface. If the applications and infrastructure supporting such services were migrated to a cloud-computing environment, the application service providers could run and maintain replicate instances of their entire application to maintain access to previous versions without the need to duplicate the hardware infrastructure underlying the

service. Alternatively, the component data and systems underlying a computational service could be archived as 'images' in the cloud in a nonactive state, which would provide an economical means of preservation for the service provider while maintaining an efficient route of access to previous versions of a computational service for investigators. This may make it more feasible for authors to cite specific versions of data, systems and services in publications describing their results, and have a means to easily direct colleagues to these resources for reproducibility and reuse.

Reproducibility through preservation

One often-overlooked barrier to reproducibility is the loss or abandonment of grant-funded databases and other computational resources¹⁷. Therefore, an advantage of cloud computing is that virtual machines and cloud-based data storage can offer a means to sustain bioinformatics projects after funding cuts, project termination or abandonment. Although it may be easy to demonstrate the need to make biomedical databases available after their original purpose has been fulfilled, from a practical standpoint, funding the maintenance of these databases has become a contentious issue for funding agencies, who must balance maintenance with declining budgets and the need to fund new investigators and initiatives. For example, the *Arabidopsis* Information Resource (TAIR), used by thousands of researchers each day, recently fell into disarray after its funding was terminated after 10 years of support from the National Science Foundation¹⁸. Instead of turning off these servers and losing access to these resources, virtual machine images of the server could have been created and distributed across the Internet and executed on-demand through cloud computing, ensuring that investments in bioinformatics resources such as TAIR could be used for years after discontinuation.

The costs associated with preserving data in a cloud-computing environment are relatively low, making it feasible for research groups and institutions to preserve data through extended gaps in funding, or to fulfill commitments to collaborators or stakeholders well beyond the defunding of the primary database infrastructure. To illustrate, the current rate for storage space in Amazon's cloud service is \$0.10 per gigabyte of storage per month, meaning that even a large 1-terabyte data set could be maintained for ~\$100 per month. As the per-gigabyte cost of storage is expected to decrease with time, it is likely that a 1-terabyte biomedical database could

be preserved in accessible form in the cloud for more than 10 years at a total cost well below \$10,000. The cost of this active preservation could be written into the budgets of all proposals for creation and renewal of biomedical databases to address the problem of preservation and access to data beyond proposed funding periods. Moreover, having a working virtual machine server along with open source code could enable distributed teams of investigators to informally support these disbanded projects.

Toward a cloud-based scientific computing commons

As nearly every scientific discipline is becoming data-driven, one of the most enticing benefits of cloud computing is the means to aggregate scientific data sets efficiently and economically. The aggregation and centralization of public scientific data in the cloud, which offers a unified, location-independent platform for data and computation, are steps toward establishing a shared 'virtual commons' for reproducible scientific computing. In this common computing environment, it would be possible to develop and implement cloud-based scientific computational analyses that retrieve public data from centralized, cloud-based master catalogs. This computational analysis could then be shared and distributed within the cloud as part of a standardized system image. Using the system image, the computational analysis could be reproducibly executed because the same source code would be executing within the same system environment, drawing the same data from the master data catalog.

Although the cloud-computing technology required to facilitate straightforward approaches to reproducible computing, such as WSSE, is now widely available, there are a number of measures that could be taken to help and encourage the utilization of cloud-based resources by the broader scientific community. Journals and funding agencies could support this vision by mandating that more types of measurements be deposited into approved cloud-based data catalogs, and funding could be provided for existing consortia to move their valuable data from isolated data silos into centralized cloud-based catalogs. As cloud computing becomes commonplace, we expect technologies to emerge that allow one to move data between cloud-computing providers, to prevent 'lock in'¹⁹. For example, the open-source Eucalyptus platform (<http://www.eucalyptus.com/>) enables one to move cloud-based resources out of the Amazon Web Services commercial platform into a private cloud infrastructure.

Table 1 Features of reproducible scientific computing in the cloud

	Traditional challenges	Cloud-computing solutions
Data sharing	<ul style="list-style-type: none"> • Large data sets difficult to share over standard internet connections; can require substantial technical resources to obtain and store. • Public data sets change frequently. Difficult to archive and share entire data repositories used for analyses. 	<ul style="list-style-type: none"> • Large data sets can be stored as 'omnipresent' resources in the cloud. Easily copied and accessed directly from any point in the cloud. • 'Snapshots' of large public data sets can be rapidly copied, archived and referenced.
Software and applications	<ul style="list-style-type: none"> • Reproducibility of results often requires replication of the precise software environment (that is, operating system, software and configuration settings) under which the original analysis was conducted. Specific versions of software or programming-language interpreters often required for reproducibility. • Analyses typically conducted by several types of software or scripts executed in a precise sequence across one or several systems as part of an analysis pipeline. Only the individual programs or scripts are usually provided with published results. Substantial technical resources typically required to recreate the pipeline used in the original analysis. • Standard software packages cannot serve all the needs of a scientific domain. Investigators develop nonstandard software and computational pipelines to facilitate computational analysis exceeding the capabilities of common tools. 	<ul style="list-style-type: none"> • Computer systems are virtualized in the cloud, allowing them to be replicated wholesale without concern for the underlying hardware. Snapshots of a fully configured system or group of systems used in analysis can be rapidly archived as digital machine images. System machine images can be copied and shared with others in the cloud, allowing reconstitution of the precise system configuration used for the original analysis. • System images can be preconfigured with common and customized software and tools in a standardized fashion to facilitate common tasks in a scientific domain (e.g., assembly of genome sequences from DNA sequencer data). Preconfigured images can be shared as public resources to promote reproducibility and follow-up studies.
System and technical	<ul style="list-style-type: none"> • Substantial computational resources might be required to replicate an analysis. Original computational analyses requiring several hundred processors to complete becoming more common. Reproducibility limited to those with requisite computational resources. • Substantial technical support often required to reproduce a computational analysis and to replicate the software and system configuration required by the analysis. Prevents reproducibility by nontechnical investigators lacking substantial IT support. 	<ul style="list-style-type: none"> • Cloud-based computational resources can be scaled up in a dynamic fashion to provide necessary computational resources. Investigators can create large computational clusters on demand and disperse upon analysis completion. • Complete digital representations of a computational pipeline can be shared as machine images along with deployment scripts that can be executed by nontechnical users to reconstitute a complete computational pipeline.
Access and preservation	<ul style="list-style-type: none"> • Grant-funded software and data repositories often disappear from the public domain after funding is discontinued or the maintainers abandon the project. Leads to loss of access by dependent users and loss of public investment into the resource. 	<ul style="list-style-type: none"> • Software, code and data from grant-funded projects can be archived and provided as publicly accessible resources in the cloud. Economies of scale in the cloud allow for active preservation of grant-funded resources for many years past funding for nominal cost. • Cloud-computing providers already show a willingness to host public scientific data sets at no cost.

Some have already called for the creation of a publicly funded cloud-computing infrastructure¹³; however, we suggest that it would be most prudent to focus funds and effort into building cloud-based scientific computing software on top of existing cloud infrastructures in such a way that they are portable across vendors. Given the economic and competitive pressures facing commercial cloud-computing vendors, as well as their substantial technical and capital resources, it is not clear that a publicly funded cloud infrastructure would offer any significant technical or economic advantages over commercial clouds. Furthermore, we suggest that, as has been observed in many aspects of computing, standards enabling cross-vendor interoperability are likely to emerge as cloud computing becomes more prominent and additional vendors increase competition in the market. Groups such as the Open Cloud Consortium (<http://opencloudconsortium.org/>) have already been formed to explore this issue. Notably, many of the popular tools for managing and interacting with cloud-based infrastructures, such as Chef (<http://www.opscode.com/chef>), are designed to be

independent of the specifics of any one cloud-computing vendor's infrastructure. Peer-to-peer data distribution technologies, such as BitTorrent which is already being leveraged in the bioinformatics community²⁰, could be used to store and distribute large biological data sets beyond the confines of any single cloud-computing provider.

In the domain of biomedicine, several efforts toward the development and distribution of cloud-based tools, systems and other resources have recently emerged²¹. Several groups have created standardized machine images with software tools and configuration settings optimized for biomedical research. The most ambitious among these so far is the J. Craig Venter Institute (Rockville, MD, USA) Cloud Bio-Linux project (<http://www.jcvi.org/cms/research/projects/jcvi-cloud-biolinux/>), which aims to provide a comprehensive and coherent system capable of a broad range of bioinformatics functionality. These preconfigured machine images are made publicly available to the research community, providing individual investigators with a standardized, tuned platform for cloud-based computational analyses. Other

efforts have gone into the creation of more comprehensive multipart systems for facilitating biocomputing in the cloud^{22,23}. A significant effort in this area is the Galaxy project²⁴, which provides a platform for large-scale genomic analysis.

Although we suggest that the WSSE approach is a pragmatic and substantial first step toward enabling reproducible scientific computing in the cloud, we acknowledge that it does not address all aspects hindering reproducibility. Foremost, software licensing constraints could prevent the use of WSSE. Many commercial software applications are restricted by licenses that constrain the number of software application instances that can run simultaneously or that restrict execution to a number of processors. These constraints might prevent an investigator from sharing some or all of the system or software components used to produce published results, limiting the effectiveness of WSSE. We also recognize that there is a need for continued research and development into reproducibility enhancements at levels above WSSE. For example, problems in data organization, systematic provenance tracking, standardization and annotation are not solved by WSSE.

Nonetheless, we suggest that WSSE can serve to initiate a value-driven movement toward general reproducible scientific data and computing into the cloud (**Table 1**), and that solutions to problems of reproducibility not solved by WSSE—many of which already exist in some form—will follow WSSE into the cloud toward the realization of a comprehensive platform for reproducible computing enabled by the technical innovations of cloud computing.

ACKNOWLEDGMENTS

J.T.D. is supported by the National Library of Medicine Biomedical Informatics Training Grant (T15 LM007033) and A.J.B. is supported by the National Institute for General Medical Sciences (R01 GM079719). We thank D. Singh for valuable discussion regarding cloud-computing technology.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

- Gentleman, R. *Stat. Appl. Genet. Mol. Biol.* **4**, Article 2 (2005).
- Gil, Y. *et al. Computer* **40**, 24–32 (2007).
- Barker, A. & van Hemert, J. in *Parallel Processing and Applied Mathematics* (eds. Wyrzykowski, R., Dongarra, J., Karczewski, K. & Wasniewski, J.) 746–753 (Springer; 2008).
- Reich, M. *et al. Nat. Genet.* **38**, 500–501 (2006).
- Hull, D. *et al. Nucleic Acids Res.* **34**, W729–W732 (2006).
- De Roure, D., Goble, C. & Stevens, R. *Future Gener. Comput. Syst.* **25**, 561–567 (2009).
- Mesirov, J.P. *Science* **327**, 415–416 (2010).
- Ball, C.A., Sherlock, G. & Brazma, A. *Nat. Biotechnol.* **22**, 1179–1183 (2004).
- Brooksbank, C. & Quackenbush, J. *OMICS* **10**, 94–99 (2006).
- Wiley, H.S. & Michaels, G.S. *Nat. Biotechnol.* **22**, 1037–1038 (2004).
- Lynch, C. *Nature* **455**, 28–29 (2008).
- Bateman, A. & Wood, M. *Bioinformatics* **25**, 1475 (2009).
- Nelson, M.R. *Science* **324**, 1656–1657 (2009).
- Osterweil, L.J., Clarke, L.A. & Ellison, A.M. *Science* **325**, 1622 (2009).
- Schwab, M., Karrenbach, M. & Claerbout, J. *Comput. Sci. Eng.* **2**, 61–67 (2000).
- Wagener, J., Spjuth, O., Willighagen, E.L. & Wikberg, J.E. *BMC Bioinformatics* **10**, 279 (2009).
- Merali, Z. & Giles, J. *Nature* **435**, 1010–1011 (2005).
- Anonymous. *Nature* **462**, 252 (2009).
- Gu, Y. & Grossman, R.L. *Eng. Sci.* **367**, 2429–2445 (2009).
- Langille, M.G. & Eisen, J.A. *PLoS ONE* **5**, e10071 (2010).
- Stein, L.D. *Genome Biol.* **11**, 207 (2010).
- Langmead, B., Schatz, M.C., Lin, J., Pop, M. & Salzberg, S.L. *Genome Biol.* **10**, R134 (2009).
- Schatz, M.C. *Bioinformatics* **25**, 1363–1369 (2009).
- Giardine, B. *et al. Genome Res.* **15**, 1451–1455 (2005).

Directed differentiation of human embryonic stem cells toward chondrocytes

Rachel A Oldershaw^{1,5}, Melissa A Baxter¹, Emma T Lowe^{1,2}, Nicola Bates¹, Lisa M Grady¹, Francesca Soncin³, Daniel R Brison^{1,4}, Timothy E Hardingham^{1,2} & Susan J Kimber¹

We report a chemically defined, efficient, scalable and reproducible protocol for differentiation of human embryonic stem cells (hESCs) toward chondrocytes. hESCs are directed through intermediate developmental stages using substrates of known matrix proteins and chemically defined media supplemented with exogenous growth factors. Gene expression analysis suggests that the hESCs progress through primitive streak or mesendoderm to mesoderm, before differentiating into a chondrocytic culture comprising cell aggregates. At this final stage, 74% (HUES1 cells) and up to 95–97% (HUES7 and HUES8 cells) express the chondrogenic transcription factor SOX9. The cell aggregates also express cell surface CD44 and aggrecan and deposit a sulfated glycosaminoglycan and cartilage-specific collagen II matrix, but show very low or no expression of genes and proteins associated with nontarget cell types. Our protocol should facilitate studies of chondrocyte differentiation and of cell replacement therapies for cartilage repair.

Articular cartilage is vitally important in the joints, providing smooth articulation and sustaining skeletal mobility. Because of its avascular nature, articular cartilage has a low intrinsic capacity for repair and is highly susceptible to damage in degenerative conditions, such as osteoarthritis. Joint degeneration with cartilage loss presents a major social and healthcare burden^{1,2}. Research on tissue engineering solutions has focused largely on mature articular chondrocytes and adult stem cells³, and less attention has been given to hESCs⁴. The development of cartilaginous tissue within hESC teratomas *in vivo* is readily demonstrated, but the proportion of chondrocytes arising from spontaneous differentiation of embryoid bodies (EBs) *in vitro* is very low^{5,6}. ESCs have been differentiated toward chondrocytes^{6–16} or multipotent mesenchymal stem-like cells¹⁷ by supplementing the culture medium with promesoderm or prochondrogenic growth factors^{6,8–12,15,18} or by co-culturing with chondrocytes^{7,9,14,16} or with developmentally immature chondroprogenitors¹³. However, the yields of chondrocyte-like cells were low, suggesting that further refinement was necessary. Some protocols reported for differentiation to other cell lineages have exploited knowledge of development, enabling step-wise differentiation of hESCs into insulin-secreting beta cells¹⁹, hepatocytes²⁰, cardiomyocytes²¹, dopaminergic neurons²² and oligodendrocytes²³. We adopted this approach for directing the differentiation of hESCs toward chondrocytes.

Musculoskeletal tissues of the limb originate from paraxial (muscle) and lateral plate mesoderm emerging from the posterior region of the primitive streak at gastrulation²⁴. Wnt3a and activin/nodal signaling function in both the formation of the primitive streak and the differentiation of

some epiblast and hESCs into a bi-potent mesendoderm population^{25–27}. The transient primitive streak and mesendoderm (primitive streak–mesendoderm) population are characterized by expression of *MIXL1* (ref. 28) and *GSC*²⁷, although genes characteristic of early mesoderm (such as *T*, also known as *BRACHYURY*²⁹) and early definitive endoderm (such as *FOXA2*; ref. 25) are also expressed. Further differentiation toward mesoderm seems to be controlled by members of the bone morphogenetic protein (BMP; in particular BMP2 and BMP4) and fibroblast growth factor (FGF) families^{30–33}. During maturation of mesoderm, expression of *T* is downregulated and cells begin to express *KDR* (also known as *FLK1* and *VEGFR2*), *PDGFRA* and *PDGFRB*^{30,31}.

The early limb bud forms as a paddle-shaped mesoblast covered by a specialized epithelium. Undifferentiated mesenchymal cells migrate into the limb field and condense to form the cartilage anlage³⁴. Central to formation of the mesenchymal condensation is expression of the high-mobility-group (HMG) domain transcription factor SOX9 (ref. 35), with accessory transcription factors SOX5 (L-SOX5) and SOX6 (ref. 36), which control chondrogenic differentiation, maintain the chondrocyte phenotype and directly regulate expression of extracellular matrix molecules, such as cartilage-specific collagen type II³⁷.

The directed differentiation protocol reported here is based on the sequence of pathways active in development, driving the differentiation of hESCs through primitive streak–mesendoderm and mesoderm intermediates to a chondrocyte population. It exploits the feeder-free, serum-free conditions we reported previously for maintenance of pluripotent hESCs³⁸.

¹North West Embryonic Stem Cell Centre, Faculty of Life Sciences, Core Technology Facility, University of Manchester, Manchester, UK. ²Wellcome Trust Centre for Cell-Matrix Research, Faculty of Life Sciences, University of Manchester, Manchester, UK. ³Faculty of Medical and Human Sciences, Core Technology Facility, University of Manchester, Manchester, UK. ⁴Department of Reproductive Medicine, Saint Mary's Hospital for Women and Children, Central Manchester University Hospitals NHS Foundation Trust, Manchester, UK. ⁵Present address: North East England Stem Cell Institute, Institute of Cellular Medicine, Newcastle University, International Centre for Life Bioscience Centre, Newcastle-Upon-Tyne, UK. Correspondence should be addressed to S.J.K. (sue.kimber@manchester.ac.uk).

Received 25 March; accepted 23 August; published online 22 October 2010; doi:10.1038/nbt.1683

RESULTS

Directed differentiation of hESCs toward chondrocytes

Our initial differentiation protocol was based on the known developmental progression of cells to primitive streak–mesendoderm and then to mesoderm. Using the HUES1 cell line, we developed a three-step directed differentiation protocol (Fig. 1 and Table 1).

Starting hESCs were maintained as feeder-free, serum-free cultures³⁸ (Fig. 2a–d), and initial differentiation, starting with cells that were 80% confluent, used a defined serum-free medium. We refined pilot protocols by testing different growth factor combinations and concentrations and varying the times of their addition and their duration (Online Methods). We assayed cells for expression of pluripotency-associated genes (Fig. 3a–d) and primitive streak–mesendodermal (Fig. 3d–f), mesodermal (Fig. 3g–k), endodermal (Fig. 3k–o), neurectodermal (Fig. 3c,p–r) and chondrocyte (Fig. 3s–x) lineage markers. This allowed us to refine the protocol to achieve more efficient differentiation. We compared our protocol to spontaneous differentiation into embryoid bodies.

Stage 1: differentiation to primitive streak–mesendoderm

Stage 1 of the protocol (days 1–3) (Table 1) aimed at differentiating pluripotent hESCs to a primitive streak–mesendoderm population based upon previous results with RPMI-base medium containing wnt3a and activin-A¹⁹. In pilot experiments, we could not maintain hESCs serum-free and feeder-free using RPMI medium. We therefore developed an alternative base medium. In this medium, the initial regime of wnt3a and activin-A¹⁹ was partly successful, but resulted in high expression of genes associated with extra-embryonic and definitive endoderm. We therefore progressively modified the protocol, reducing the initial activin-A from 50 ng ml⁻¹ to 25 ng ml⁻¹ at day 2 and 10 ng ml⁻¹ at day 3. Wnt3a (25 ng ml⁻¹) was also extended to day 3, with FGF2 (20 ng ml⁻¹) from day 2 onwards and BMP4 (40 ng ml⁻¹) added on day 3.

During stage 1, the number of cells approximately doubled (Supplementary Fig. 1). Cells retained hESC morphology, during stage 1 (Fig. 2e,f) with large nuclei and prominent nucleoli (Fig. 2e,f). The pluripotency genes *POU5F1* (also known as *OCT4*) and *NANOG* continued to be expressed at levels similar to those of hESC cultures

(Fig. 3a,b), whereas an apparent decrease in *SOX2* was not statistically significant (Fig. 3c). However, there was evidence of differentiation toward a mesendodermal population by the end of stage 1. The cell adhesion molecule E-cadherin is expressed by pluripotent hESCs³⁹ and is also associated with a mesendoderm phenotype²⁷; its expression of the E-cadherin gene, *CDH1* (also known as *ECAD*), was accordingly upregulated by 1.8-fold ($P < 0.01$) in stage 1 (Fig. 3d). Similarly, expression of *GSC* (encoding gooseoid) increased 3.3-fold ($P < 0.05$) (Fig. 3e), and expression of *T*, a gene expressed in early mesoderm⁴⁰, but also shown to be expressed by murine ES cells with potential for both mesoderm and endoderm⁴¹, rapidly increased 800-fold ($P < 0.001$) (Fig. 3g). Immunofluorescence at stage 1 showed that Brachyury and gooseoid were expressed by >over 95% of the cells. Consistent with observations in pluripotent hESCs, 95% of cells expressed E-cadherin (as assessed by flow cytometry), and this proportion was unchanged through stage 1. Immunolocalization showed E-cadherin to be expressed more strongly in some regions of the culture (Supplementary Fig. 2). At stage 1, we also detected a five-fold increase in *GATA4* ($P < 0.05$) (Fig. 3I), typically associated with commitment to definitive endoderm. These data together suggest that stage 1 yielded a cell population enriched for mesendoderm.

Stage 2: differentiation to mesoderm

During stage 2 of the protocol (days 4–8) (Table 1), BMP4 and FGF2 supplementation continued, wnt3a and activin-A were removed, and the expression of endoderm-specifying genes was reduced further by including follistatin (100 ng ml⁻¹). In stage 2, neurotrophin-4 (NT4) (2 ng ml⁻¹) was added to promote cell survival³⁸. During this stage, cultures were expanded by two passages (Table 1), reflecting a high proliferation rate (Fig. 2g) (Supplementary Fig. 1). At the end of stage 2 (Fig. 2g,h), the cells were 95% confluent and contained phase-bright cell clusters 40–80 μm in diameter (circled in Fig. 2g,h) at a density of 60 ± 5.5 clusters per mm² (mean ± s.e.m.; $n = 4$). These stained with safranin O, showing the accumulation of sulfated glycosaminoglycan (sGAG) (Fig. 4a,b). Comparing gene expression of cells at the end of stage 2 with hESCs, expression of *OCT4* and *SOX2* decreased to 0.1% ($P < 0.001$) and 3.3% ($P < 0.01$), respectively; expression of *CDH1* decreased to 6% ($P < 0.01$) and of *GSC* to 8% ($P < 0.05$).

MIXL1 expression, which is often used to identify primitive streak–mesendodermal cells, was upregulated later than predicted, increasing 6.3-fold by the end of stage 2, compared with hESCs ($P < 0.05$) (Fig. 3f). There was evidence of differentiation to an intermediate mesoderm cell population at the end of stage 2: *KDR*, expressed by multipotent mesoderm³⁰, was upregulated by 5.5-fold ($P < 0.05$) between stages 1 and 2 (Fig. 3j), and *CXCR4*, which is used to identify mesodermal populations destined to form hemangioblastic cells⁴¹, was also upregulated by 1.5-fold ($P < 0.05$) over the same period (Fig. 3k). Notably, *SOX9*, which is expressed in all chondrogenic cells³⁵, but also in early development in other lineages, was upregulated fivefold ($P < 0.05$) at the end of stage 2 (Fig. 3s), and there was much lower expression of definitive endoderm lineage genes. *GATA4*, which had been high at stage 1, declined to just 0.5% of that level by stage 2

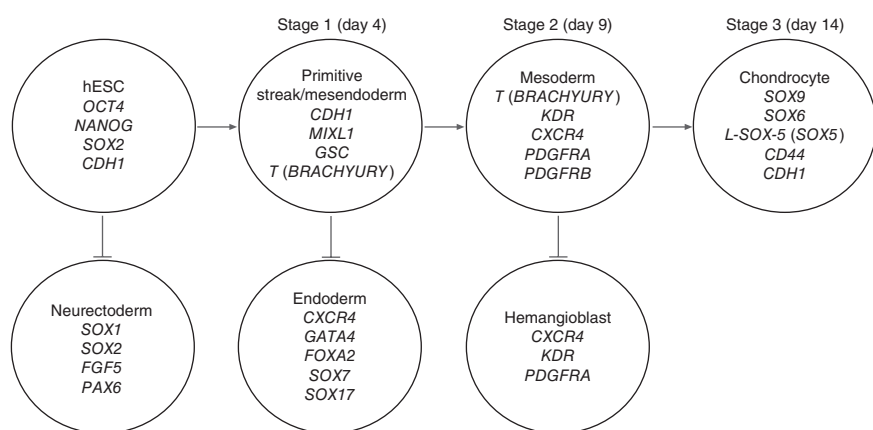


Figure 1 Schematic of directed differentiation protocol in three stages. In stage 1, pluripotent hESCs are directed toward a primitive streak–mesendoderm population; in stage 2, differentiation proceeds to a mesoderm population; and in stage 3, toward chondrocytes. As some genes are expressed in different cell lineages and at different stages, the developmental status of each cell population was characterized by expression of panels of marker genes including *SOX2*, which is expressed by both pluripotent hESCs and cells derived from the neurectoderm germ layer, *CDH1*, expressed on pluripotent and mesendoderm cells and *CXCR4*, used to identify cell lineages from both the endoderm and mesoderm-derived hemangioblast.

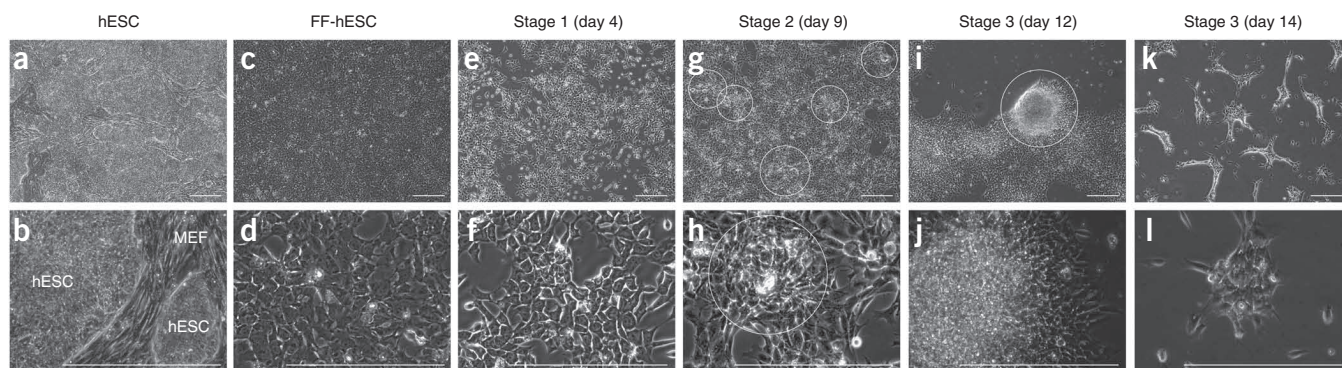


Figure 2 Morphology of hESC cultures (HUES1) at different stages of the protocol. **(a,b)** Pluripotent hESCs on a mouse embryonic fibroblast (MEF) feeder layer. Cell cultures were heterogeneous with hESCs forming individual, tightly packed colonies. **(c,d)** Pluripotent hESCs cultured on a fibronectin matrix in a defined medium. hESCs appeared as a homogeneous 2D monolayer, with individual cells appearing larger than those in colonies maintained on feeder layers. The cells had characteristic hESC morphology, with a high nucleus-to-cytoplasm ratio and prominent nucleoli. FF-hESC; feeder-free hESC. **(e,f)** At the end of stage 1, cells were ~80% confluent and still retained morphological features of pluripotent hESCs. **(g,h)** At stage 2, cell cultures were densely packed, with phase-bright cell clusters distributed throughout the culture (circled). These clusters formed 3D nodules with cells clustered into a 'rosette-like' morphology (circled). **(i,j)** During stage 3 (day 12), the flatter cells surrounding the 3D cell aggregates began to detach, leaving behind the 3D cell aggregates (circled). Cultures are on fibronectin-gelatin substrate and were imaged immediately before passaging onto gelatin as detailed in **Table 1**. **(k,l)** By end of stage 3 (day 14), intermediate cells between the cell aggregates had also become detached. All scale bars, 100 μm .

($P < 0.05$) (**Fig. 3l**), and the expression of *FOXA2* and *SOX17* was significantly lower (**Fig. 3m,n**). To test for cells destined for other lineages, we checked for genes expressed much later in endoderm differentiation. We did not detect expression of *HNF1A*, *HNF4A*, *PROX1*, *ALB* or *AFP* at any stage of our protocol (**Supplementary Fig. 3**), confirming that the increase in *SOX9* expression was not associated with differentiation to pancreatic or hepatocyte lineages.

Stage 3: differentiation to chondrocytes

Stage 3 of the protocol (days 9–14; **Table 1**) included a graded switch from BMP4 to GDF5, which was increased to 40 ng ml^{-1} by day 11. By day 12 (before passage), the cell clusters formed during stage 2 had increased in size (circled in **Fig. 2i,j**), with less firmly attached, flatter cells between the clusters. At the end of stage 3, the cell density was lower than earlier (**Fig. 2k,l**), although there was an 8.5-fold ($P < 0.05$) increase in cell numbers over the entire differentiation protocol (**Supplementary Fig. 1**). The flatter cells were no longer present, leaving many small independent aggregates of cells with rounded chondrocyte-like morphology. Formation of these aggregates

appeared similar to the initial stages of chondrogenesis³⁴, and the aggregates were of similar size, but they were more compact, and occurred at slightly lower density than the cell clusters at stage 2 (46 ± 2.1 aggregates per mm^2 ($n = 4$)). They also stained more strongly for safranin O (**Fig. 4c,d**), and the staining was sensitive to chondroitinase ABC; this confirmed that the cell aggregates accumulated chondroitin sulfate proteoglycans within the matrix (**Fig. 4e–h**). We quantified the accumulation of sGAG (**Fig. 4i**), revealing a 78-fold increase in sGAG production per cell between stage 2 and 3 cultures ($P < 0.05$).

These cells did not express *OCT4*, *NANOG*, *SOX2* and *CDH1*, indicating loss of pluripotency, and we did not detect expression of genes associated with intermediate cell populations, including *GSC*, *MIXL1* and *T*. Expression of *KDR* declined to 0.005% ($P < 0.05$) of the values observed in stage 2, suggesting that the protocol directed differentiation away from a hemangioblastic population. To our surprise, there was no change in expression of *PDGFRA*, reportedly associated with mesodermal cells^{27,42}, throughout the protocol (**Fig. 3h**), although expression of its homolog, *PDGFRB*^{27,43}, had increased 5.5-fold by the end of stage 3 ($P < 0.01$) (**Fig. 3i**).

Table 1 Protocol for stages 1–3 of the differentiation regime for hESCs

Stage	Day	Culture area (cm^2)	Plating density (cells cm^{-2})	Matrix substrate	WNT3A (ng ml^{-1})	Activin-A (ng ml^{-1})	FGF2 (ng ml^{-1})	BMP4 (ng ml^{-1})	Follistatin (ng ml^{-1})	GDF5 (ng ml^{-1})	NT4 (ng ml^{-1})
1	1	9.62	100,000	FN	25	50					
1	2	9.62		FN	25	25	20				
1	3	9.62		FN	25	10	20	40			
2	4	9.62		FN			20	40	100		2
2	5	48.11	46,500	FN			20	40	100		2
2	6	48.11		FN			20	40	100		2
2	7	48.11		FN			20	40	100		2
2	8	192.44	46,500	FN:gel			20	40			2
3	9	192.44		FN:gel			20	20		20	2
3	10	192.44		FN:gel			20	20		20	2
3	11	192.44		FN:gel			20			40	2
3	12	288.66	34,000	Gel			20			40	2
3	13	288.66		Gel			20			40	2
3	14										

HESCs were fed daily in base medium supplemented with growth factors at the stated concentrations. Differentiating cultures were expanded by passage with trypsin. Passage and cell expansion is shown as the increase in culture area, based on starting with a single 35-mm dish. Tissue culture dishes were coated with fibronectin, fibronectin/gelatin mixed at a ratio of 50:50 or gelatin. FN, fibronectin; gel, gelatin; NT4, neurotrophin-4.

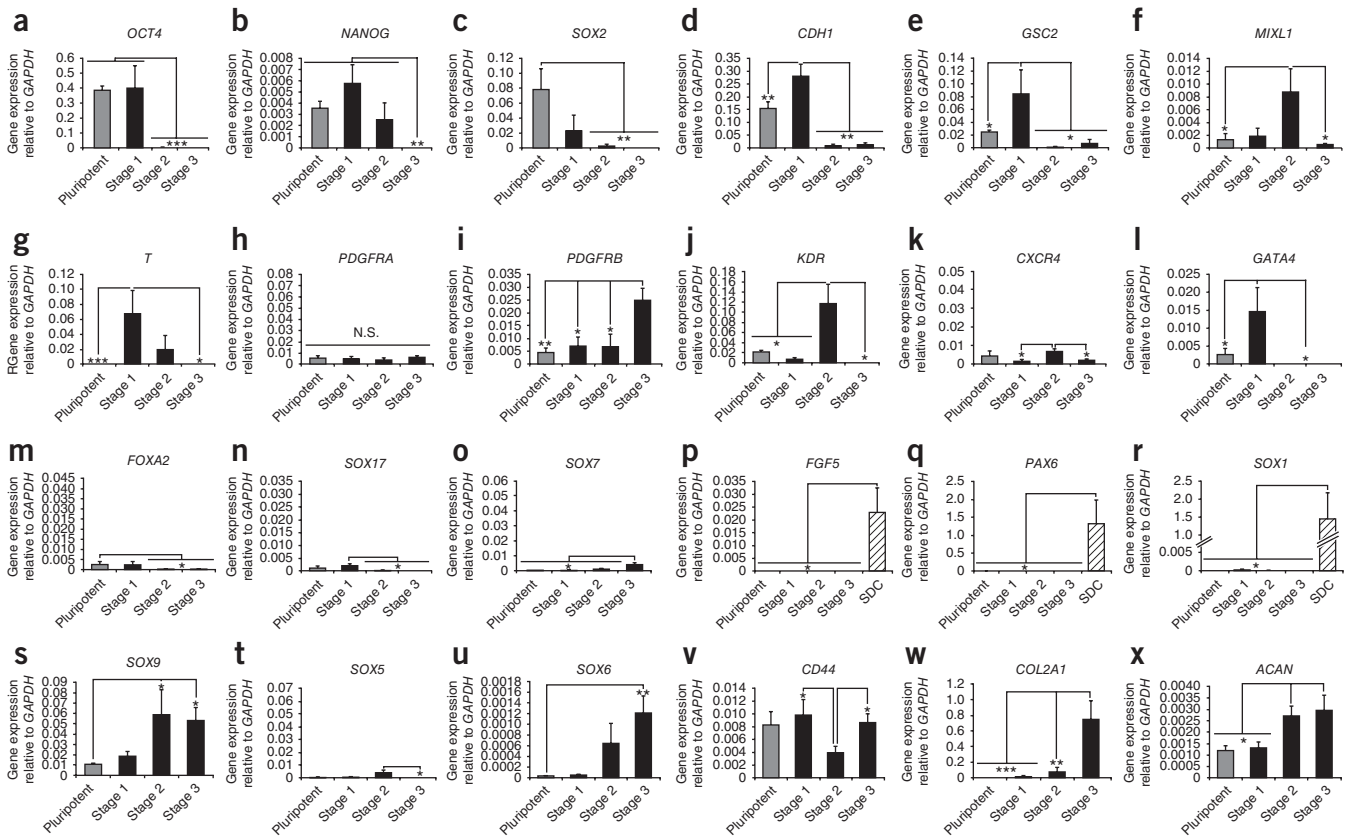


Figure 3 Gene expression analysis of hESCs at different stages of the protocol. Pluripotent hESCs (HUES1; gray bars) and differentiating cultures at the end of stage 1 (day 4), stage 2 (day 9) and stage 3 (day 14) (black bars) were analyzed for gene expression, as denoted in **Figure 1**. (a–d) Genes associated with hESC pluripotency. (d–f) Genes expressed by cell lineages from the primitive streak–mesendoderm. (g–k) Genes expressed by mesodermal cell types. (k–o) Genes expressed by endoderm cell types. (p–r) Genes expressed by neurectodermal cell types (also in c). (s–x) Genes expressed by chondrocytes. hESCs transiently expressed genes associated with a primitive streak–mesendoderm phenotype, before expression of genes associated with mesodermal cell lineages. Embryoid body–derived spontaneous differentiation cultures (SDC; hatched bars) taken at day 14 were used as a positive control to confirm the specificity of the primers used and to verify the pluripotent phenotype of the original hESC cultures. Gene expression was normalized to GAPDH. Values represent means \pm s.e.m. ($n = 4$). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

The presence of chondroblastic cells at the end of stage 3 was assessed on the basis of the expression of *SOX9*, *SOX5* (*L-SOX5*), *SOX6*, *CD44*, *COL2A1* and *ACAN* (**Fig. 3s–x**). *SOX9* expression was similar to that in stage 2, and there was a 3.7-fold upregulation of its transcription co-factor, *SOX6* ($P < 0.01$). Although we recorded some modest upregulation of *SOX5* at the end of stage 2, it was inconsistent, and expression fell by half at the end of stage 3 ($P < 0.05$). *CD44* expression was maintained during stage 1, but was subsequently lost, before returning to the level observed in hESCs. *COL2A1*, a downstream gene target of *SOX9* (ref. 37) that encodes a key cartilage matrix protein, showed a 370-fold increase in expression between hESCs and cells at the end of stage 3 ($P < 0.01$). *ACAN*, which encodes another cartilage extracellular matrix component, was also upregulated 2.5-fold between hESCs and cells at the end of stage 3 ($P < 0.05$). We did not detect *COL10A1*, which is expressed by terminally differentiating, hypertrophic chondrocytes, in our directed differentiation protocol; this contrasts with CD105-positive adult mesenchymal stem cells undergoing chondrogenic differentiation (**Supplementary Fig. 4**).

To further understand the phenotype of the cells generated by this protocol, we also investigated the regulation of genes expressed by related mesodermal cell lineages (**Supplementary Fig. 5**). *RUNX2* (also known as *CBFA1*), which encodes a transcription factor

expressed during osteoblast differentiation⁴⁴, was upregulated tenfold between stages 1 and 2 ($P < 0.01$) before decreasing by 50% at the end of stage 3 ($P < 0.05$). *PPARG* γ , which regulates adipogenesis⁴⁵, and *SCLERAXIS*, a transcription factor expressed during tenogenic differentiation⁴⁶, were expressed at a low level throughout directed differentiation. These data suggest that our directed differentiation protocol is specific for differentiation toward chondrocytes.

Characterization of stage 3 cells

We immunostained stage 3 cultures for differentiation markers (**Fig. 5** and **Supplementary Fig. 6**) and compared them to spontaneously differentiating cultures (**Supplementary Fig. 7**) from embryoid bodies. The latter expressed proteins associated with all three germ layers as well as developmental intermediates. In contrast, stage 3 cultures were more homogeneous, with little nontarget marker expression (**Supplementary Fig. 6**). *OCT4* and *NANOG* expression were not detected by immunostaining at the end of stage 3. Expression of *SOX2* protein was weak and detected solely within the cell cytoplasm, providing further evidence of loss of pluripotency. The presence of developmentally immature cells was assessed by immunostaining for *MIXL1*, *Brachyury* and *SOX17* (**Supplementary Fig. 6**), all of which gave only faint signals, restricted to the cytoplasm, at

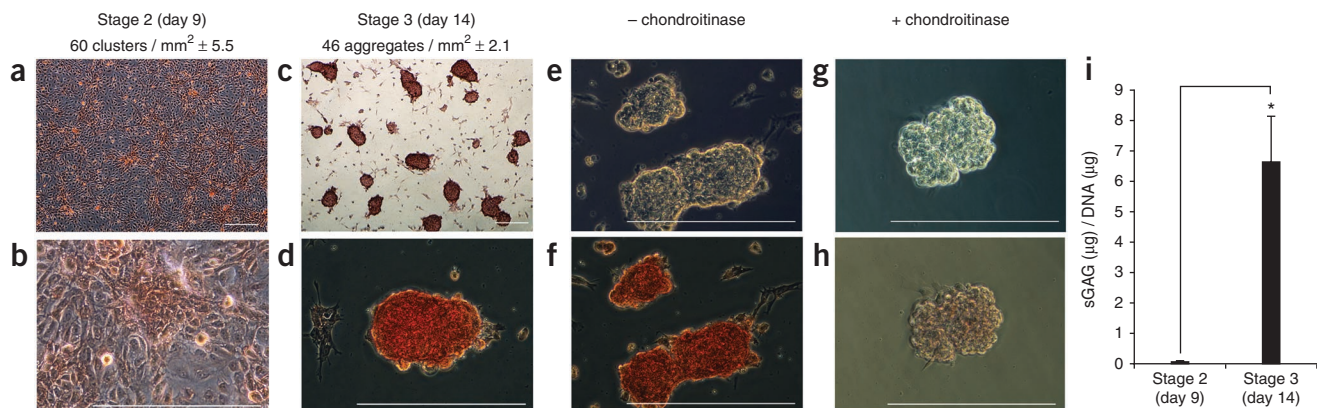


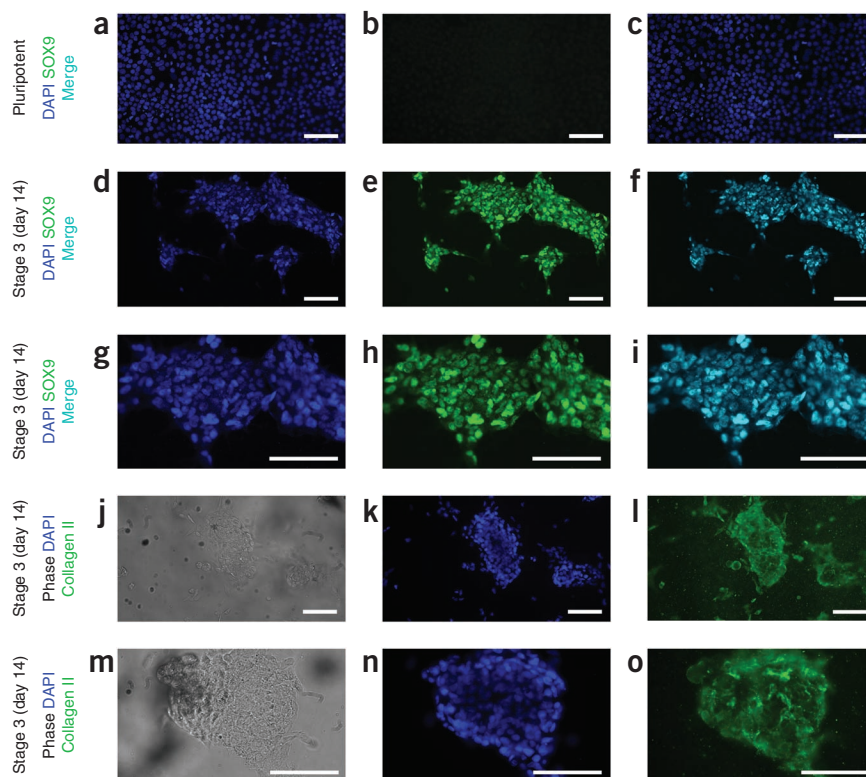
Figure 4 Sulfated glycosaminoglycan accumulation during directed differentiation of hESCs (HUES1) to chondrocytes. (a,b) Cell clusters that formed throughout stage 2 cultures stained discretely with safranin O at day 9, indicating the accumulation of sGAG. (c,d) At stage 3, independent cell aggregates were larger than the cell clusters in stage 2 and at day 14 showed markedly stronger staining for safranin O. The frequency of cell clusters staining for safranin O in stage 2 cultures and cell aggregates in stage 3 cultures was determined as described in Online Methods. Values represent means \pm s.e.m. ($n = 4$). (e–h) To demonstrate the specificity of safranin O staining for sGAG, some cultures were treated with chondroitinase ABC before staining. Control stage 3 cultures before (e) and after (f) staining with safranin O. Stage 3 cultures treated with chondroitinase ABC before (g) and after (h) staining with safranin O. All scale bars, 100 μ m. (i) Quantification of sGAG production per cell at stage 2 and 3. Values represent means \pm s.e.m. ($n = 4$). * $P < 0.05$.

the end of stage 3, suggesting that most cells had differentiated beyond these intermediate stages. This contrasted with spontaneous embryoid body differentiation, which yielded a more heterogeneous cell population (Supplementary Fig. 7). Similarly, immunostaining for GATA4 and SOX7 (definitive and visceral endoderm, respectively) showed markedly reduced abundance of both proteins, and no nuclear localization, by the end of stage 3 (Supplementary Fig. 6). Among mesodermal markers, PDGFR α was absent at the end of stage 3, and KDR (FLK1) was considerably reduced, but PDGFR β was expressed (Supplementary Fig. 6). Immunostaining for SOX9, CD44 and collagen II (Fig. 5 and Supplementary Fig. 6) showed that at the end of stage 3, cells had a chondrocyte-like phenotype. Although we detected basal SOX9 gene expression in pluripotent hESCs and at the end of stage 1 (Fig. 3s), we detected no SOX9 protein in hESCs (Fig. 5a–c). In contrast, at the end of stage 3, SOX9 was highly expressed, and most SOX9-positive cells showed nuclear localization (Fig. 5d–i). CD44 was detected on stage 3 cells as punctate immunostaining at the cell surface (Supplementary

Fig. 6). Collagen type II was also abundant, providing a clear indication that the cells had developed a chondrocyte-like phenotype (Fig. 5j–o).

By flow cytometry, we found that SOX9 protein was expressed by $74.8\% \pm 3.1\%$ ($n = 4$) of the cells at the end of stage 3 (Fig. 6a,b) although the proportion of cells expressing the hyaluronan receptor CD44 ($34.9\% \pm 2.1\%$; $n = 4$) was much lower (Fig. 6c,d). We also analyzed expression of CD105, which is characteristic of undifferentiated mesenchymal stem cells (Supplementary Fig. 4). CD105 was expressed on only $12.7\% \pm 2.5\%$ ($n = 4$) of the cells at the end of the protocol (Fig. 6e,f).

Figure 5 Immunofluorescence of SOX9 and collagen II. Proteins were indirectly labeled with secondary Alexa Fluor 488 antibodies (green channel) and cell nuclei labeled with DAPI (blue channel). (a–c) Expression of chondrogenic transcription factor SOX9 was low in pluripotent hESC cultures (HUES1). (d–i) At the end of the differentiation protocol (day 14), SOX9 was highly expressed and SOX9 protein was localized within the nuclei of aggregated cells. (j–o) Phase images and immunofluorescence of collagen II deposited within the matrix of cell aggregates at the end of the differentiation protocol. Scale bar, 100 μ m.



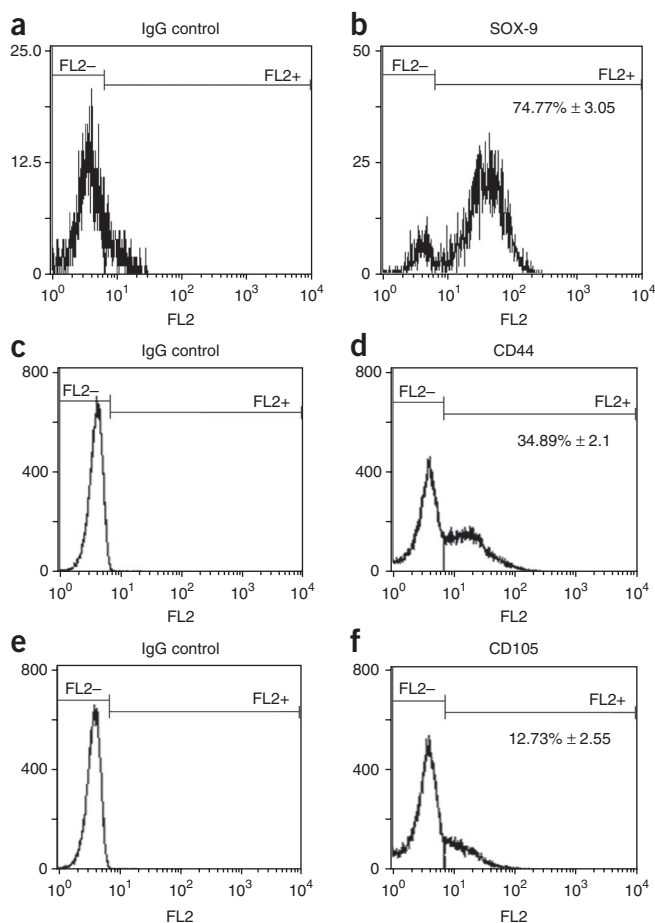


Figure 6 Flow cytometry analyses of HUES1-derived cells at the end of stage 3 (day 14). (a–f) Cells were analyzed for expression of the chondrocyte transcription factor SOX9 (a,b), the cell surface receptor CD44 (c,d) and the adult stem cell surface antigen CD105 (e,f). Immunological control (a) for SOX-9. The proportion of SOX9-expressing cells (b) was $74.8\% \pm 3.1$. Immunological control (c) for CD44. The proportion of CD44-expressing cells (d) was $34.1\% \pm 2.1$. Immunological control (e) for CD105. The proportion of CD105-expressing cells (f) was $12.7\% \pm 2.6$. Values represent means \pm s.e.m. ($n = 4$) FL-2 indicates exclusion gate for flow cytometry analysis.

an endochondral ossification route⁴⁷. Together with other age- and disease-related limitations of bone marrow stem cells^{48,49}, this finding suggests that embryonic stem cells may provide a more appropriate source of clinically relevant chondrogenic cells.

We started the protocol with a feeder-free system, which eliminates undefined bio-active molecules secreted by feeders and creates a more homogeneous culture. We postulate that this allows more uniform exposure to growth factors, nutrients and oxygen tension, all of which influence hESC fate decisions⁵⁰. The two-dimensional (2D) culture format appears to offer more potential than 3D embryoid bodies for producing a high-yield, scalable protocol. The protocol achieves an 8.5-fold increase in cell number (from days 0 to 12).

Maintenance of cell viability proved difficult in the absence of feeders and serum-rich media. We found that a proprietary anti-oxidant supplement (B27) together with insulin, transferrin and selenium (ITS) supported cell metabolism and enhanced survival. Although this base medium conferred some advantages, it did not support prolonged culture on its own. From day 4 of the protocol, we added a low concentration of NT4 to maintain hESC survival³⁸. Furthermore, it is likely that FGF2 added from day 2 also contributed to cell survival³⁸. In the initial step, cells were cultured on fibronectin as done for hESC maintenance³⁸. As culture proceeded, fibronectin became dispensable and we could progressively passage cells on gelatin. As differentiation proceeded, the cells may have assembled a more complex extracellular matrix, which supported cell survival.

In assessing gene expression, we artificially designated the end of each stage as a single time point in a continuous differentiation program. However, the peaks in transient expression differ between genes and may occur either before or after times of analysis. In stage 1, hESCs were directed to a primitive streak–mesendoderm intermediate population characterized by upregulation of E-cadherin, goosecoid and brachyury. Notably, the transient upregulation of the primitive streak–mesendoderm marker *MIXL1* was highest at the end of the stage 2 when *GSC*, *CDH1* and *FOXA2* had all been downregulated. It is likely that *MIXL1* transcription was suppressed in the presence of goosecoid⁵¹ and so increased only after the latter had decreased, reflecting expression during *in vivo* development.

We consider that optimization of stage 1, priming the cells for later signaling cues, was the most crucial step in developing an efficient differentiation protocol, because as in preliminary experiments it had the most influence over the final cellular phenotype. The concentration of activin was particularly important. When we used high concentrations of activin, substantial expression of endoderm-associated genes, specifically, *SOX17*, *FOXA2* and *GATA4*, could still be observed. This was true even when the stage 1 medium included insulin, an inhibitor of definitive endoderm differentiation⁵², and the stage 2 and 3 media included the promesoderm-prochondrogenic cytokines BMP4 and GDF5. This potent effect of activin on cell lineage specification *in vitro* has been described¹⁹ and may reflect the morphogenic gradients along the anterior-posterior axis of the primitive streak, with activin concentration highest where endoderm is specified

The rounded morphology and, expression of SOX9, CD44 and collagen II, together with lack of CD105, indicate that the protocol produces cells more developmentally advanced toward chondrocytes than are mesenchymal stem cells.

We developed the protocol throughout with HUES1 cells, but we also tested it with HUES7 and HUES8, which showed very similar cell morphology and characteristics to HUES1 during the protocol. Notably, they both yielded a higher proportion of SOX9-positive cells at the end of the protocol (stage 3) than did HUES1 cells (HUES7, $95\% \pm 1.01\%$, $n = 4$; HUES8, $97\% \pm 0.3\%$, $n = 4$; **Supplementary Fig. 8**).

DISCUSSION

We have developed a chemically defined protocol for the directed differentiation of hESCs toward chondrocytes. We divide the protocol into three stages to incorporate the transient enrichment of intermediate mesodermal cell populations known to precede chondrogenesis during development. This three-stage protocol is highly efficient, resulting in chondrogenic cells of which, in three different hESC lines, 74–97% express the chondrocyte-associated transcription factor SOX9. During the differentiation process, the cells acquired a rounded chondrocyte-like morphology, expressed the hyaluronan cell surface receptor CD44 and formed three-dimensional (3D) aggregates in which collagen type II and a sGAG-rich matrix were deposited. Notably, we detected no expression of *COL10A1* (associated with hypertrophy) at any stage during the protocol; this contrasts with chondrogenic differentiation of bone marrow stem cells, which yields cells in which *COL10A1* is upregulated, suggestive of a hypertrophic phenotype and thus, indicative of differentiation along

anteriorly²⁶. Cells from the posterior of the primitive streak, which differentiate into mesoderm, are exposed to lower concentrations of activin. Conversely, low concentrations of activin at the beginning of a pancreatic beta cell differentiation protocol reduced the proportion of insulin-expressing cells¹⁹. Following the developmental approach of that protocol, we applied wnt3a for 3 d and added BMP4 on day 3 of stage 1, as these cytokines are prominent in the posterior primitive streak and direct cells toward mesoderm³³.

In stage 2 of the protocol, the population resembling primitive streak–mesoderm was further differentiated by continued treatment with BMP4, which is involved particularly in posterior mesoderm patterning³³. Inhibition of FGF signaling attenuates BMP4-mediated mesoderm specification and anterior-posterior mesoderm patterning³², so we continued FGF2, but removed activin to increase the efficiency of differentiation to mesoderm. This was confirmed by the upregulation of *PDGFRB*, *KDR* and *CXCR4* (refs. 30,31,33) in stage 2. *FOXA2*, *SOX17* and *GATA4*, associated with definitive endoderm, were all downregulated at the end of stage 2, supporting mesodermal differentiation, without any expansion of endoderm-lineage cells. Whether the latter cells die, or are redirected, cannot be ascertained from this study.

BMPs, particularly BMP4, not only mediate mesoderm specification but also instruct uncommitted mesenchymal cells to become chondroprogenitors⁵³. During stage 2, cells aggregated into 3D clusters reminiscent of the mesenchymal condensations that precede chondrogenesis. Indeed, genes associated with chondrogenesis, *SOX9* and its downstream target collagen II, increased at the end of stage 2.

In stage 3, we aimed to enrich the cells for a chondrocyte-like population. Although treatment with BMP4 is chondrogenic and can initiate the formation of mesenchymal condensations⁵³, the effect of BMP4 alone also caused, here and elsewhere, the expression of the hemangioblast markers *KDR* and *CXCR4* (ref. 31). We therefore added a related member of the TGF β superfamily, GDF5 (ref. 53), which has a role in chondrogenesis and maintenance of a mature chondrogenic phenotype⁵⁴ and is reportedly more potent than BMP4 at initiating mesenchymal condensations and cartilage nodule formation⁵³.

Analysis of the 3D aggregates at the end of stage 3 showed that cells had assembled a collagen type II and sGAG-rich matrix, characteristic of a cartilaginous tissue². Aggregates expressed *SOX9*, prominently localized to the cell nucleus, an indication of functional activity. Flow cytometry showed that 74% of the final population in HUES1 and 95–97% in HUES 7 and HUES8 expressed *SOX9*. This may reflect inherent differences between genetically distinct cell lines in their tendency to form chondrogenic cells, as previously reported for hESCs⁵⁵, or to respond to the *in vitro* conditions. It is possible that more than 95% of all three HUES lines might express *SOX9* given sufficient time in culture. Although cells at stage 3 aggregated and acquired a rounded morphology, some remained adhered to the plastic. Chondrocyte adhesion to tissue culture plastic induces actin stress fiber formation and loss of *SOX9* expression⁵⁶. This effect may variably influence the number of cells expressing *SOX9*. The proportion in the final population of HUES1 cells expressing the hyaluronan cell surface receptor CD44 (34%) was less than the proportion expressing *SOX9*. In mature cartilage, CD44 provides a link between the chondrocyte and the pericellular matrix through its interaction with hyaluronan. CD44 expression is responsive to the amount of matrix surrounding the cell⁵⁷, and the lower proportion of cells expressing CD44 may reflect the immaturity of the extracellular matrix within the aggregates. We therefore suggest that these cells are immature chondrocytes.

At the end of stage 3, the cells did not express the pluripotency markers *OCT4* and *NANOG*, an important safety consideration for clinical translation of the protocol, as undifferentiated hESCs may give rise to teratomas. In contrast to undirected differentiation, our protocol produced few nontarget cell types.

The success of the protocol probably results from a combination of directing hESCs through each developmental stage and providing a culture environment in which target cells thrive and nontarget cells are lost. Derivation of chondrocytic cells from hESCs is very efficient, as shown with three genetically different hESC lines. Moreover, the system is fully chemically defined and scalable. These advantages suggest that the protocol is appropriate for the study of chondrogenesis during mammalian development and as a basis for producing chondrocytes suitable for clinical translation.

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 was funded by the North West Regional Development Agency; North West Embryonic Stem Cell Centre is also supported by the UK Medical Research Council and the UK National Institute for Health Research Biomedical Research Funding Scheme. We thank N. Hanley (University of Manchester) for the gift of human fetal cDNA.

AUTHOR CONTRIBUTIONS

R.A.O., D.R.B., T.E.H. and S.J.K. were responsible for study concept and design, analysis and interpretation of data and preparation of the manuscript. R.A.O., M.A.B., E.T.L., N.B., F.S. and L.M.G. were responsible for the acquisition of data.

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. Goldring, M.B. & Goldring, S.R. Osteoarthritis. *J. Cell. Physiol.* **213**, 626–634 (2007).
2. Hardingham, T.E. Articular cartilage. in *Oxford Textbook of Rheumatology* (eds. Maddison, P.J., Isenberg, D.A., Woo, P. & Glass, D.N.) 325–334, (Oxford University Press, Oxford, UK, 2004).
3. Hardingham, T.E., Oldershaw, R.A. & Tew, S.R. Cartilage, *SOX9* and Notch signals in chondrogenesis. *J. Anat.* **209**, 469–480 (2006).
4. Murry, C.E. & Keller, G. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell* **132**, 661–680 (2008).
5. De Sousa, P.A. *et al.* Clinically failed eggs as a source of normal human embryo stem cells. *Stem Cell Res.* **2**, 188–197 (2009).
6. Kawaguchi, J., Mee, P.J. & Smith, A.G. Osteogenic and chondrogenic differentiation of embryonic stem cells in response to specific growth factors. *Bone* **36**, 758–769 (2005).
7. Bigdell, N. *et al.* Coculture of human embryonic stem cells and human articular chondrocytes results in significantly altered phenotype and improved chondrogenic differentiation. *Stem Cells* **27**, 1812–1821 (2009).
8. Boyd, N.L., Robbins, K.R., Dhara, S.K., West, F.D. & Stice, S.L. Human embryonic stem cell-derived mesoderm-like epithelium transitions to mesenchymal progenitor cells. *Tissue Eng. Part A* **15**, 1897–1907 (2009).
9. Hoben, G.M., Willard, V.P. & Athanasiou, K.A. Fibrochondrogenesis of hESCs: growth factor combinations and cocultures. *Stem Cells Dev.* **18**, 283–292 (2009).
10. Koay, E.J., Hoben, G.M. & Athanasiou, K.A. Tissue engineering with chondrogenically differentiated human embryonic stem cells. *Stem Cells* **25**, 2183–2190 (2007).
11. Kramer, J. *et al.* Embryonic stem cell-derived chondrogenic differentiation *in vitro*: activation by BMP-2 and BMP-4. *Mech. Dev.* **92**, 193–205 (2000).
12. Lee, E.J. *et al.* Novel embryoid body-based method to derive mesenchymal stem cells from human embryonic stem cells. *Tissue Eng. Part A* **16**, 705–715 (2009).
13. Sui, Y., Clarke, T. & Khillan, J.S. Limb bud progenitor cells induce differentiation of pluripotent embryonic stem cells into chondrogenic lineage. *Differentiation* **71**, 578–585 (2003).
14. Vats, A. *et al.* Chondrogenic differentiation of human embryonic stem cells: the effect of the micro-environment. *Tissue Eng.* **12**, 1687–1697 (2006).

15. Yang, Z., Sui, L., Toh, W.S., Lee, E.H. & Cao, T. Stage-dependent effect of TGF- β 1 on chondrogenic differentiation of human embryonic stem cells. *Stem Cells Dev.* **18**, 929–940 (2009).
16. zur Nieden, N.I., Kempka, G., Rancourt, D.E. & Ahr, H.J. Induction of chondro-, osteo- and adipogenesis in embryonic stem cells by bone morphogenetic protein-2: effect of cofactors on differentiating lineages. *BMC Dev. Biol.* **5**, 1 (2005).
17. Lian, Q. *et al.* Derivation of clinically compliant MSCs from CD105⁺, CD24⁻ differentiated human ESCs. *Stem Cells* **25**, 425–436 (2007).
18. Nakagawa, T., Lee, S.Y. & Reddi, A.H. Induction of chondrogenesis from human embryonic stem cells without embryoid body formation by bone morphogenetic protein 7 and transforming growth factor β 1. *Arthritis Rheum.* **60**, 3686–3692 (2009).
19. D'Amour, K.A. *et al.* Production of pancreatic hormone-expressing endocrine cells from human embryonic stem cells. *Nat. Biotechnol.* **24**, 1392–1401 (2006).
20. Hay, D.C. *et al.* Efficient differentiation of hepatocytes from human embryonic stem cells exhibiting markers recapitulating liver development in vivo. *Stem Cells* **26**, 894–902 (2008).
21. Laflamme, M.A. *et al.* Cardiomyocytes derived from human embryonic stem cells in pro-survival factors enhance function of infarcted rat hearts. *Nat. Biotechnol.* **25**, 1015–1024 (2007).
22. Yan, Y. *et al.* Directed differentiation of dopaminergic neuronal subtypes from human embryonic stem cells. *Stem Cells* **23**, 781–790 (2005).
23. Nistor, G.I., Totoiu, M.O., Haque, N., Carpenter, M.K. & Keirstead, H.S. Human embryonic stem cells differentiate into oligodendrocytes in high purity and myelinate after spinal cord transplantation. *Glia* **49**, 385–396 (2005).
24. Winslow, B.B., Takimoto-Kimura, R. & Burke, A.C. Global patterning of the vertebrate mesoderm. *Dev. Dyn.* **236**, 2371–2381 (2007).
25. Gadue, P., Huber, T.L., Paddison, P.J. & Keller, G.M. Wnt and TGF- β signaling are required for the induction of an *in vitro* model of primitive streak formation using embryonic stem cells. *Proc. Natl. Acad. Sci. USA* **103**, 16806–16811 (2006).
26. Sumi, T., Tsuneyoshi, N., Nakatsuji, N. & Suemori, H. Defining early lineage specification of human embryonic stem cells by the orchestrated balance of canonical Wnt/ β -catenin, Activin/Nodal and BMP signaling. *Development* **135**, 2969–2979 (2008).
27. Tada, S. *et al.* Characterization of mesendoderm: a diverging point of the definitive endoderm and mesoderm in embryonic stem cell differentiation culture. *Development* **132**, 4363–4374 (2005).
28. Izumi, N., Era, T., Akimaru, H., Yasunaga, M. & Nishikawa, S. Dissecting the molecular hierarchy for mesendoderm differentiation through a combination of embryonic stem cell culture and RNA interference. *Stem Cells* **25**, 1664–1674 (2007).
29. Wilkinson, D.G., Bhatt, S. & Herrmann, B.G. Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature* **343**, 657–659 (1990).
30. Ema, M., Takahashi, S. & Rossant, J. Deletion of the selection cassette, but not *cis*-acting elements, in targeted Flk1-lacZ allele reveals Flk1 expression in multipotent mesodermal progenitors. *Blood* **107**, 111–117 (2006).
31. Era, T. *et al.* Multiple mesoderm subsets give rise to endothelial cells, whereas hematopoietic cells are differentiated only from a restricted subset in embryonic stem cell differentiation culture. *Stem Cells* **26**, 401–411 (2008).
32. Faloon, P. *et al.* Basic fibroblast growth factor positively regulates hematopoietic development. *Development* **127**, 1931–1941 (2000).
33. Zhang, P. *et al.* Short-term BMP-4 treatment initiates mesoderm induction in human embryonic stem cells. *Blood* **111**, 1933–1941 (2008).
34. Hall, B.K. & Miyake, T. All for one and one for all: condensations and the initiation of skeletal development. *Bioessays* **22**, 138–147 (2000).
35. Akiyama, H., Chaboissier, M.C., Martin, J.F., Schedl, A. & de Crombrugge, B. The transcription factor Sox9 has essential roles in successive steps of the chondrocyte differentiation pathway and is required for expression of Sox5 and Sox6. *Genes Dev.* **16**, 2813–2828 (2002).
36. Lefebvre, V., Behringer, R.R. & de Crombrugge, B. L-Sox5, Sox6 and Sox9 control essential steps of the chondrocyte differentiation pathway. *Osteoarthritis Cartilage* **9** Suppl. A, S69–S75 (2001).
37. Lefebvre, V., Huang, W., Harley, V.R., Goodfellow, P.N. & de Crombrugge, B. SOX9 is a potent activator of the chondrocyte-specific enhancer of the pro α 1(II) collagen gene. *Mol. Cell. Biol.* **17**, 2336–2346 (1997).
38. Baxter, M.A. *et al.* Analysis of the distinct functions of growth factors and tissue culture substrates necessary for the long-term self-renewal of human embryonic stem cell lines. *Stem Cell Res.* **3**, 28–38 (2009).
39. Eastham, A.M. *et al.* Epithelial-mesenchymal transition events during human embryonic stem cell differentiation. *Cancer Res.* **67**, 11254–11262 (2007).
40. Kispert, A., Herrmann, B.G., Leptin, M. & Reuter, R. Homologs of the mouse *Brachyury* gene are involved in the specification of posterior terminal structures in *Drosophila*, *Tribolium*, and *Locusta*. *Genes Dev.* **8**, 2137–2150 (1994).
41. Kubo, A. *et al.* Development of definitive endoderm from embryonic stem cells in culture. *Development* **131**, 1651–1662 (2004).
42. Takenaga, M., Fukumoto, M. & Hori, Y. Regulated nodal signaling promotes differentiation of the definitive endoderm and mesoderm from ES cells. *J. Cell Sci.* **120**, 2078–2090 (2007).
43. Betsholtz, C., Karlsson, L. & Lindahl, P. Developmental roles of platelet-derived growth factors. *Bioessays* **23**, 494–507 (2001).
44. Ducy, P. Cbfa1: a molecular switch in osteoblast biology. *Dev. Dyn.* **219**, 461–471 (2000).
45. Rosen, E.D. The transcriptional basis of adipocyte development. *Prostaglandins Leukot. Essent. Fatty Acids* **73**, 31–34 (2005).
46. Schweitzer, R. *et al.* Analysis of the tendon cell fate using Scleraxis, a specific marker for tendons and ligaments. *Development* **128**, 3855–3866 (2001).
47. Pelttari, K. *et al.* Premature induction of hypertrophy during *in vitro* chondrogenesis of human mesenchymal stem cells correlates with calcification and vascular invasion after ectopic transplantation in SCID mice. *Arthritis Rheum.* **54**, 3254–3266 (2006).
48. Coipeau, P. *et al.* Impaired differentiation potential of human trabecular bone mesenchymal stromal cells from elderly patients. *Cytotherapy* **11**, 584–594 (2009).
49. Murphy, J.M. *et al.* Reduced chondrogenic and adipogenic activity of mesenchymal stem cells from patients with advanced osteoarthritis. *Arthritis Rheum.* **46**, 704–713 (2002).
50. Sachlos, E. & Auguste, D.T. Embryoid body morphology influences diffusive transport of inductive biochemicals: a strategy for stem cell differentiation. *Biomaterials* **29**, 4471–4480 (2008).
51. Izzi, L. *et al.* Foxh1 recruits Gsc to negatively regulate Mixl1 expression during early mouse development. *EMBO J.* **26**, 3132–3143 (2007).
52. McLean, A.B. *et al.* Activin efficiently specifies definitive endoderm from human embryonic stem cells only when phosphatidylinositol 3-kinase signaling is suppressed. *Stem Cells* **25**, 29–38 (2007).
53. Hatakeyama, Y., Tuan, R.S. & Shum, L. Distinct functions of BMP4 and GDF5 in the regulation of chondrogenesis. *J. Cell. Biochem.* **91**, 1204–1217 (2004).
54. Pacifici, M., Koyama, E. & Iwamoto, M. Mechanisms of synovial joint and articular cartilage formation: recent advances, but many lingering mysteries. *Birth Defects Res. C Embryo Today* **75**, 237–248 (2005).
55. Osafune, K. *et al.* Marked differences in differentiation propensity among human embryonic stem cell lines. *Nat. Biotechnol.* **26**, 313–315 (2008).
56. Tew, S.R. & Hardingham, T.E. Regulation of SOX9 mRNA in human articular chondrocytes involving p38 MAPK activation and mRNA stabilization. *J. Biol. Chem.* **281**, 39471–39479 (2006).
57. Grover, J. & Roughley, P.J. Expression of cell-surface proteoglycan mRNA by human articular chondrocytes. *Biochem. J.* **309**, 963–968 (1995).

ONLINE METHODS

Cell culture. *2D monolayer culture of hESCs.* All cell culture media and supplements were purchased from Gibco, Invitrogen, unless otherwise stated. hESC lines HUES1, HUES7 and HUES8 (passage 24–42; <http://www.mcb.harvard.edu/melton/hues/>) were cultured for bulk up on mitomycin C-inactivated MF1 × CD1 mouse embryonic fibroblasts (MEFs) in hESC medium (Knockout DMEM, 20% (vol/vol) Knockout Serum Replacement, 1% (vol/vol) ITS supplement, 1% (vol/vol), non-essential amino acids, 2 mM L-glutamine, 50 U ml⁻¹ penicillin, 50 µg ml⁻¹ streptomycin, 90 µM β-mercaptoethanol supplemented with 10 ng ml⁻¹ FGF2 (Biosource, Invitrogen)). Feeder-free hESC culture was carried out as described³⁸. Briefly, hESCs were lifted from MEF feeders with trypsin (PAA Laboratories), centrifuged at 720g for 3 min before being resuspended in feeder-free culture medium (DMEM:F12, 0.1% (wt/vol) BSA (Sigma), 2 mM L-glutamine, 1% (vol/vol) nonessential amino acids, 2% (vol/vol) B27, 1% (vol/vol) N2 liquid supplement, 90 µM β-mercaptoethanol, 10 ng ml⁻¹ activin-A (Peprotech), 40 ng ml⁻¹ FGF2, 2 ng ml⁻¹ neurotrophin-4 (Peprotech)). hESCs were replated onto fibronectin-coated (50 µg ml⁻¹; Millipore) tissue culture plastic. Initially, feeder-free cultures were established by subculturing confluent hESC cultures at a ratio of 1:1 to achieve ~90% plating efficiency. Once established, confluent feeder-free cultures were subcultured by trypsin passage and replated onto fresh fibronectin substrate at a cell density of 5 × 10⁴ cells cm⁻², a ratio of ~1:4. Feeder-free hESCs were cultured over at least three passages to ensure the removal of all MEF cells.

Directed differentiation studies. hESCs were cultured on tissue culture plastic, coated with appropriate matrix substrates in a basal medium (DMEM:F12, 2 mM L-glutamine, 1% (vol/vol) ITS, 1% (vol/vol) nonessential amino acids, 2% (vol/vol) B27, 90 µM β-mercaptoethanol) supplemented with appropriate growth factors. The differentiation protocol described and characterized in Results was developed through empirical and systematic optimization of each stage of differentiation. Preliminary protocols were carried out on the embryonic stem cell line HUES7 with directed differentiation driven by culturing the cells in base medium supplemented with combinations of exogenous growth factors. Differentiation to our target chondrogenic cells was assessed by semiquantitative PCR (sqPCR) for expression of target and nontarget genes associated with primitive streak–mesodermal (MIXL1), mesodermal (T, PDGFRβ, KDR), endodermal (SOX17, FOXA2, GATA4, AFP), neur ectodermal (SOX1, PAX6, FGF5) and chondrogenic (SOX9, CD44, collagen II). To assess the progression and efficiency of our directed differentiation over random spontaneous differentiation, sqPCR was carried out on cDNA generated from outgrowths of embryoid bodies which had been cultured over equivalent lengths of time. Progressive refinement of our pilot protocols was carried out by modifying the growth factors applied, the concentrations at which they were added and the kinetics of application. The criterion for successful refinement of our protocol was the elimination of expression of genes associated with non-target cell lineages while establishing and maintaining maximal expression of genes associated with chondrogenic differentiation. The ranges of each growth factor concentration and days of application assessed are listed below.

Wnt3a was assessed at 25 ng ml⁻¹ on day 1 alone, days 1 and 2 inclusive and days 1–3 inclusive. Activin-A was assessed at 100–10 ng ml⁻¹ over days 1–8. BMP4 was assessed at 10–40 ng ml⁻¹ over days 3–10. NT4 was assessed at 0–2 ng ml⁻¹ added over days 4–14. Follistatin was assessed at 50–100 ng ml⁻¹ over days 4–7. FGF2 was assessed at 10–20 ng ml⁻¹ over days 2–14. GDF5 was assessed at 10–40 ng ml⁻¹ over days 9–14.

Optimal concentrations and temporal application of growth factors determined by these preliminary experiments are presented as our directed differentiation protocol (see Results and **Table 1**; Wnt3a, R&D Systems), human follistatin 300 (Sigma), BMP4, GDF5 (Peprotech). Cell counts were carried out at each passaging event to quantify the expansion of the cell culture during directed differentiation. For immunostaining at the end of stage 3, cells were subcultured on day 12 onto glass coverslips coated with gelatin.

Embryoid body formation and culture. Feeder-free HUES1 cells were detached from tissue culture plastic by incubation with collagenase IV for 10 min at 37 °C, 5% CO₂. Small clumps of cells were resuspended in DMEM (DMEM, 10%

(vol/vol) FBS, 110 µg ml⁻¹ sodium pyruvate, 100 U ml⁻¹ penicillin, 100 µg ml⁻¹ streptomycin) and plated as a suspension culture in bacteriological grade Petri dishes. After 7 d, embryoid bodies were dissociated by trypsin digestion and plated on either FBS-coated tissue culture plastic or glass coverslips (for immunostaining) for the required number of days.

Culture and chondrogenic differentiation of human mesenchymal stem cells. Human mesenchymal stem cells were derived from human bone marrow mononuclear cell preparations (Lonza) and cultured as described⁵⁸. At passage 3, the cells were formed into cell aggregates (500,000 cells per aggregate) and cultured for up to 14 d in chondrogenic medium: high-glucose DMEM, 2 mM L-glutamine, 110 µg ml⁻¹ sodium pyruvate, 100 U ml⁻¹ penicillin, 100 µg ml⁻¹ streptomycin, 50 µg ml⁻¹ ascorbic acid 2-phosphate, 40 µg ml⁻¹ L-proline, 100 nM dexamethasone, 10 ng ml⁻¹ TGFβ-3, 1% (vol/vol) ITS+1 supplement. Cells were centrifuged into aggregate cultures at day 0.

Gene expression analysis. *RNA purification.* Total RNA was extracted from monolayer cell cultures using Qiashredder and RNeasy mini kits (Qiagen) according to the manufacturer's instructions. Total RNA was extracted using Tri-reagent (Sigma) from cell aggregates ground using Molecular Grinding Resin (Geno Technology) according to the manufacturer's instructions. All RNA was treated with 10 U µl⁻¹ RNA sample DNase I (Invitrogen). Total RNA (1 µg per 25-µl reaction) was reverse transcribed using 200 U Moloney murine leukemia virus (M-MLV) reverse transcriptase primed with 0.5 µg random hexamer oligonucleotides (Promega).

Quantitative PCR analysis. PCR was carried out on a MJ Research Opticon 2 using Sybr Green I detection (Eurogentec) according to the manufacturer's instructions. Gene specific primers were designed using Applied Biosystems Primer Express software and used at a final concentration of 300 nM. Relative expression levels were normalized to GAPDH and calculated using the 2^{-ΔCt} method as described⁵⁸. All primers were purchased from Invitrogen and are described in **Supplementary Table 1**.

Immunofluorescence, immunohistochemistry and histological analysis. For sGAG detection, fixed cells were stained in 0.1% (wt/vol) safranin O (in 0.1% (vol/vol) in acetic acid) for 5 min at room temperature. For determining the specificity of safranin O for sGAG, some cultures were incubated with 1 unit ml⁻¹ chondroitinase ABC (Sigma) in chondroitinase buffer (50 mM Tris, pH 8.0), 60 mM sodium acetate, 0.02% (wt/vol) BSA) for 30 min at 37 °C before safranin O staining. For determining the frequency of cell clusters staining for safranin O at stage 2 and cell aggregates at stage 3, cultures were imaged on an Olympus CKX41 microscope with Olympus U-CMAD3 camera. Cell clusters/aggregates which contained >12 cells were counted over predetermined independent fields of view (0.35 mm² (700 µm × 500 µm)) chosen with-out preference. The average of ten fields was taken for each of four replicates.

Immunofluorescence. For immunolocalization, samples were blocked (0.1% (wt/vol) BSA, 10% (vol/vol) serum of animal in which secondary was raised, in PBS) for 1 h, incubated at 4 °C for 16 h with primary antibody (**Supplementary Table 2**) or irrelevant primary and for 1 h in appropriate secondary antibody (donkey anti-mouse IgG Alexa 488 or donkey anti-goat IgG Alexa 488) both diluted (1:250) in blocking buffer. Cells were counterstained with 300 nM 4'-6-diamidino-2-phenylindole (DAPI) for 5 min before mounting with Vectorshield (Vector Laboratories). For immunofluorescence of collagen II protein, cell cultures were treated with chondroitinase ABC before blocking. Appropriate immunological controls were carried out in parallel. Antibodies were purchased as described in **Supplementary Table 2**.

Biochemical analysis. *Preparation of samples.* Cell cultures were digested in 10 U ml⁻¹ papain in papain buffer (0.1 M sodium acetate, 2.4 mM EDTA, 5 mM L-cysteine, pH 5.8) (all from Sigma) in 50 µl cm⁻² of culture area for 2 h at 60 °C.

Quantification of DNA. The amount of DNA in papain-digested samples was determined using a Quant-iT PicoGreen ds DNA kit (Invitrogen) according to the manufacturer's instructions.

1,9-dimethylmethylene blue (DMMB) assay for sGAG. DMMB solution (100 μl of 16 mg l^{-1} 1,9-dimethylmethylene blue, 40 mM glycine, 40 mM NaCl, 9.5 mM HCl, pH 3.0) was added to papain-digested samples (40 μl) and the absorbance (A_{540}) of the resultant solution read immediately. The amount of sGAG in each sample was calculated against a known standard of shark chondroitin sulfate (Sigma).

Flow cytometry. Cells were detached and digested to a single-cell suspension using trypsin. For labeling of intracellular antigens, cells were fixed in ice-cold methanol for 10 min at $-20\text{ }^{\circ}\text{C}$ and further permeabilized by incubation with 1% (wt/vol) BSA, 0.5% (vol/vol) triton-X-100 in PBS for 15 min. Cells were incubated with primary antibody (goat anti-human SOX-9, $8\text{ }\mu\text{g ml}^{-1}$) diluted in ice-cold buffer (10% (vol/vol) FBS, 1% (wt/vol) sodium azide in PBS) overnight at $4\text{ }^{\circ}\text{C}$. For labeling of cell surface antigens, cells were incubated with primary antibody (mouse anti-human CD44

$50\text{ }\mu\text{g ml}^{-1}$), goat anti-human Endoglin/CD105 ($5\text{ }\mu\text{g ml}^{-1}$) in ice-cold buffer (2% (vol/vol) FBS, 0.1% (wt/vol) sodium azide in PBS) for 2 h. Cells were incubated with appropriate secondary antibodies (donkey anti-mouse IgG Alexa 488 or donkey anti-goat IgG Alexa 488) diluted to 1:250. Cell labeling was analyzed using a Cell Lab Quanta SC MPL flow cytometer with software version 1.0 (Beckman Coulter).

Statistical analysis. Data sets were analyzed for normal distribution using the one-sample Kolmogorov-Smirnov test. Normally distributed data were analyzed by parametric two-sample *t*-test. Data which were not normally distributed were analyzed by an equivalent nonparametric test.

58. Oldershaw, R.A. *et al.* Notch signaling through Jagged-1 is necessary to initiate chondrogenesis in human bone marrow stromal cells but must be switched off to complete chondrogenesis. *Stem Cells* **26**, 666–674 (2008).

Spatially addressed combinatorial protein libraries for recombinant antibody discovery and optimization

Hongyuan Mao¹, James J Graziano¹, Tyson M A Chase¹, Cornelia A Bentley¹, Omar A Bazirgan¹, Neil P Reddy¹, Byeong Doo Song^{1,3} & Vaughn V Smider^{1,2}

Antibody discovery typically uses hybridoma- or display-based selection approaches, which lack the advantages of directly screening spatially addressed compound libraries as in small-molecule discovery. Here we apply the latter strategy to antibody discovery, using a library of ~10,000 human germline antibody Fabs created by *de novo* DNA synthesis and automated protein expression and purification. In multiplexed screening assays, we obtained specific hits against seven of nine antigens. Using sequence-activity relationships and iterative mutagenesis, we optimized the binding affinities of two hits to the low nanomolar range. The matured Fabs showed full and partial antagonism activities in cell-based assays. Thus, protein drug leads can be discovered using surprisingly small libraries of proteins with known sequences, questioning the requirement for billions of members in an antibody discovery library. This methodology also provides sequence, expression and specificity information at the first step of the discovery process, and could enable novel antibody discovery in functional screens.

Recombinant antibodies are now a major class of therapeutics with over 20 US Food and Drug Administration (FDA)-approved molecules in use for a myriad of diseases. Currently only two main approaches exist for antibody discovery: (i) animal immunization followed by hybridoma¹ or B-cell isolation^{2,3} and (ii) library display technologies. Immunization led to the initial discovery of all currently marketed therapeutic antibodies; however, it is a time consuming process that usually leads to antibodies that require further engineering like humanization and affinity maturation. Molecular display libraries have contributed greatly to antibody discovery and engineering, with high affinity and specificity being generated through numerous mutagenesis and selection techniques⁴. Highly diverse libraries containing antibody V_H and V_L genes linked to their encoded proteins can be produced in a number of formats including phage^{5,6}, ribosome⁷, yeast⁸ and mammalian cell^{9,10} display systems. Such libraries can often reach complexities of 10¹⁰–10¹⁴, which allow rare high-affinity binders to be identified through competitive selection on target antigens⁴.

Display-based libraries have limitations because they rely on competitive selection based on target affinity¹¹. They may miss protein molecules with properties that are often incompatible with high affinity, such as agonists, partial agonists and antagonists, and modulators of target function. Modulating a pathway through ‘rheostat’-based therapeutics might be pharmacologically more desirable than turning a pathway ‘on’ or ‘off’ with high-affinity binders. Many small molecules with modulatory properties are currently approved drugs and are highly sought after¹². Another drawback of display-based libraries is that several major pharmacologic target classes (e.g., G protein-coupled receptors and ion channels) are largely refractory

to antibody discovery because of either their low expression in cell lines or difficulties in producing properly folded purified proteins for affinity selections.

To enable discovery of novel antibodies without the drawbacks of display systems, we sought to exploit the common key property of other methods for interrogating molecular interactions, such as microarrays^{13,14} and combinatorial chemistry libraries^{15,16}. Namely, each library member is spatially addressed, wherein each molecule is at a discrete location with a priori knowledge of the contents at each position. This format allows each molecule to be interrogated independently, providing data on every library member simultaneously. Moreover, it provides control of reactant concentrations in biochemical assays, allows multiplexed analysis and is also amenable to functional types of cell-based assays. Thus, at a fundamental level, antibody discovery currently relies on genetic selection mechanisms, which do not offer the advantages of the one-by-one screening approaches used to discover small molecules. Technical challenges and costs, though, make it impractical to express and purify the individual proteins needed to generate a spatially addressed antibody library that contains on the order of 10⁸ members, the minimum required size suggested by some empirical and theoretical work¹⁷.

A number of lines of evidence, however, suggest that a spatially addressed library comprising tens of thousands of individual members should be sufficient for successful low-affinity lead discovery. First, the actual diversity of naive B cells capable of responding to antigen at any given time in a mouse is only 10⁴, based on experimental data¹⁸ and paratope minimal repertoire theory¹⁹. Second, transgenic mice with a severely restricted V_H combinatorial repertoire can produce antibodies against most antigens, calling into

¹Fabrus LLC, La Jolla, California, USA. ²The Scripps Research Institute, La Jolla, California, USA. ³Present address: Scripps Korea Antibody Institute, Chuncheon-si, Korea. Correspondence should be addressed to V.V.S. (vsmider@scripps.edu).

Received 24 June; accepted 27 September; published online 24 October 2010; doi:10.1038/nbt.1694

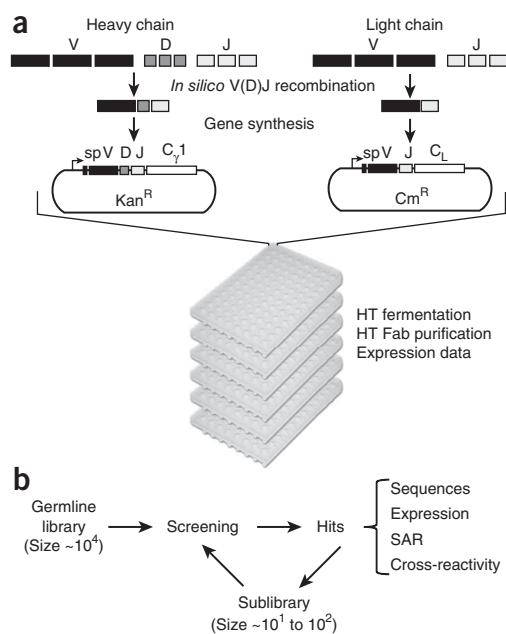


Figure 1 A spatially addressed antibody library for discovery and optimization. (a) Schematic of spatially addressed combinatorial library design and generation. (b) Schematic of screening and optimization.

question the importance of large numbers of antigen receptors *in vivo*²⁰. Third, germline antibodies contain flexible complementarity-determining region (CDR) loops, which gives them the ability to adopt more than one conformation of the antigen binding site^{21–26}, suggesting that a functional single precursor could respond to more than one antigen^{23,25,27,28}. Fourth, as all of the antibody V (variable), D (diversity) and J (joining) gene segments are now known²⁹, along with the machinery and mechanisms that generate recombinational and junctional diversity³⁰, it is possible to calculate that all rearranged VDJ heavy chains, when combined with all VJ rearranged light chains, could produce a theoretical repertoire of only $\sim 2 \times 10^7$ in the absence of junctional diversity. This number can be further reduced by eliminating redundant amino acid sequences and highly homologous V regions. Although junctional sequence diversification through insertions and deletions can greatly expand the diversity of the repertoire, terminal deoxynucleotidyl transferase (TdT) is not active in the neonatal or B1a B-cell repertoires, and TdT knockout mice show no obvious humoral defects³¹. Furthermore, a protein array screen of >27,000 human proteins against 12 naive antibodies identified four specific interactions³², suggesting the feasibility of a screening approach with small libraries.

Based on the above considerations, we developed an approach for antibody discovery and optimization using a spatially addressed library of 10⁴ lead molecules (Fig. 1). Arrayed libraries of highly purified soluble human germline Fabs in microtiter plates were produced using *de novo* DNA synthesis and automated parallel protein expression and purification. Each well contained a unique VDJ recombined heavy chain and VJ recombined light chain, with sequences known a priori. These protein-based libraries are analogous to combinatorial chemistry small-molecule libraries. A high-throughput binding assay using multiplex electrochemiluminescence (ECL) was developed, and antibodies with micromolar binding affinity were identified for several targets. Two hits were further affinity matured. Using the sequence-activity-relationship

(SAR) information from adjacent sequence space of either the heavy chain or light chain, we quickly identified important binding regions on the Fabs. This information provided guidance for further optimization toward higher binding affinity. The matured antibodies had both partial and full antagonist activities. We validated the function of these affinity matured antibodies in cell-based assays. Hence, we have demonstrated that a small-molecule discovery paradigm can be applied to large-molecule protein therapeutics, opening up new opportunities in biologic discovery.

RESULTS

Library generation

Our goal was to design a library with the highest structural diversity in a relatively small number of Fabs that could be screened in a spatially addressed format. Initially, all human germline V, D and J sequences were gathered. We chose to omit redundant V-region sequences (>90% homology) and also restricted the D-region sequence reading frames to those encoding the most hydrophilic sequences, which are favored *in vivo*³³. All V(D)J recombinations were then generated *in silico* (without nucleotide insertions or deletions) resulting in a database of 3,022,272 (7,128 heavy chains \times 424 light chains) sequences. A reverse BLAST algorithm was applied in an effort to assess the diversity of all recombined sequences. Following this, 353 of the most diverse sequences were chosen for the heavy chain and 60 for the light chain, creating a library of 21,120 possible Fab sequences. In this scheme, a diverse structural repertoire was anticipated because most framework regions were used. A more focused subset was also created where the commonly used V_H3-23 region³⁴ was recombined with all possible D and J segments. Thus, the complete germline VDJ diversity of 690 recombinants was sampled for this single heavy-chain framework.

We created a Fab library by the scalable combinatorial association of separate heavy and light chain genes (Fig. 1a). Plasmid vectors containing compatible replication origins³⁵ were designed to facilitate the separate cloning of variable regions of heavy or light chains (V_H and V_L) in frame with the appropriate constant regions³⁶. With $\sim 1,000$ heavy chain clones and 60 light chain clones, we could theoretically create a Fab library size of 60,000. In practice, typically 96 heavy-chain genes were individually co-transformed with a single light-chain gene. This was done in parallel with six light-chain genes per iteration. Co-transformed cultures (6 \times 96-well plates) served as the inocula for a production run in the Piccolo automated protein expression and purification system³⁷ followed by a secondary purification on protein G running in parallel. Each cycle took ~ 7 d and theoretically would result in 576 unique highly purified Fabs. However, $\sim 10\%$ were very low expressers with yields ranging from several micrograms to ~ 400 μ g per sample, largely depending on the heavy- and light-chain sequence composition (Supplementary Figs. 1 and 2).

Library screening and hit identification

As the number of germline Fabs in our library accumulated into the thousands, we designed a screening assay platform meeting several predefined criteria. We required that the assay: (i) be robust enough to detect weak binders with micromolar range affinities, (ii) be amenable to multiplex analysis to provide binding information on several targets of interest simultaneously and (iii) use low volumes to minimize the consumption of the library. Toward these goals, we developed a binding assay using Meso Scale Discovery's multiplex ECL platform, where each well of a 96-well plate contains nine targets and one blank (Fig. 2a). The screening was carried out in a homogenous binding environment, where Fab and a ruthenium-labeled secondary

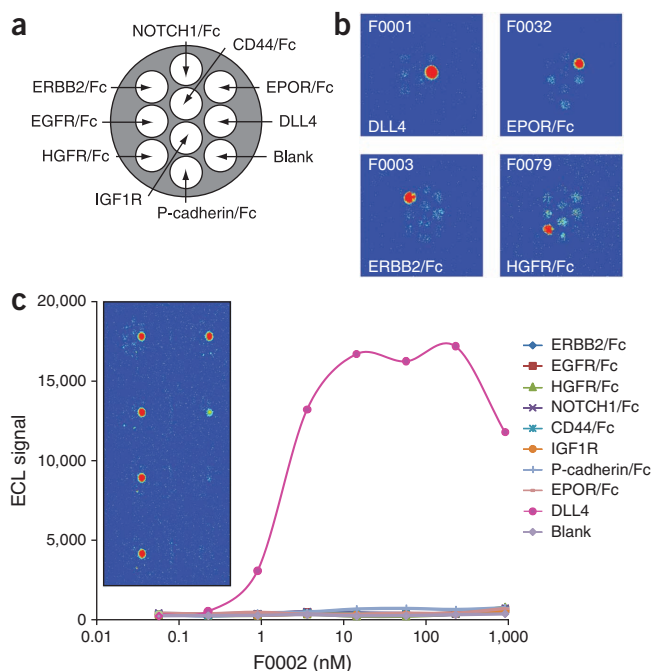


Figure 2 Screening and identifying hits using ECL detection. **(a)** Antigen map of the ten-spot 96-well plate used in the screen. **(b)** Examples of four hits identified from ECL detection: F0001 against DLL4, F0003 against ERBB2/Fc, F0032 against EPOR/Fc and F0079 against HGFR/Fc. Quantified ECL signals are listed in **Supplementary Table 1**. **(c)** Titration of F0002 confirms its binding affinity and specificity to DLL4. Serial dilutions of F0002 were tested (ECL image is inset), and the resulting signal for each of the ten spots is graphed as a function of F0002 concentration.

and anti-cancer stem cell effects in animal models^{44–46}, presumably by promoting excessive nonproductive neovascularization in the tumor. High-affinity antagonist antibodies targeting DLL4 are now in clinical trials, however, their safety has been questioned owing to hepatic toxicity with repeated use⁴⁷. Thus, alternative activities in targeting DLL4 could be clinically relevant.

Upon screening the library, we identified 18 hits against human DLL4. The V_H and V_L compositions of these hits cover several different families, suggesting diverse DLL4 binding epitopes. The binding constants (K_d) of four Fabs ranged from 0.5 μ M (F0030) to 38 μ M (F0003) using SPR. Two Fabs (F0001, 4.8 μ M; and F0002, 0.73 μ M) were chosen for further affinity maturation because they have completely different V_H and V_L sequences—one with a λ -light chain, the other with a κ -light chain, respectively.

Affinity maturation

To avoid creating a large mutagenesis library and to avoid applying display-based selection, we developed a systematic process to optimize the low-affinity binders identified from the germline library. In our spatially addressed format, we can compare binding affinities of the hit versus nearby non-hit Fabs in sequence space because their sequence identities are known a priori. Such comparisons can reveal SARs of important CDR(s) and potentially important residues within the CDRs for binding the target.

In the case of F0002, a closely related non-hit Fab was first identified (**Fig. 3a**). Both Fabs have the same light chain and share the same heavy chain V segment. Only six amino acids in the heavy chain CDR3 are different, accounting for their differences in binding to DLL4. We applied alanine-scanning mutagenesis within the CDR3 and observed that substitution of E100, Y101, S105, E107 or Q110 caused a reduction in DLL4 binding, whereas mutation of S102, S103, S104 or H111 either improved or did not affect binding affinity to DLL4 (**Supplementary Table 4**). Because these latter four are adjacent to residues important for binding, we concluded that these wild-type residues may not be optimal and should be further substituted with other amino acids. From a mutant library, we identified that the S102A (M0007), S103P (M0006), S104F (M0003) and H111F single mutations each enhanced DLL4 binding (**Table 1**). The combined S102A, S103P and S104F triple mutant (M0008) led to sixfold improvement in affinity, and addition of H111F as a quadruple mutant (M0010) conferred a tenfold improvement compared to the parent F0002 (**Table 1** and **Fig. 3b,c**). By applying the same mutation principles sequentially, we identified four additional mutants (I51V, N52L, S54T and G56H) in the heavy-chain CDR2, three mutants (S28N, S30D and S31H) in CDR1 and two mutants (S52L and A55S) in CDR2 of the light chain that contribute to binding improvements (**Table 1** and **Fig. 3b,c**). By combining all these mutations we effectively matured the DLL4 binding affinity from 730 nM (F0002) to 1.7 nM (M0026), a 430-fold enhancement (**Table 1** and **Fig. 3b,c**).

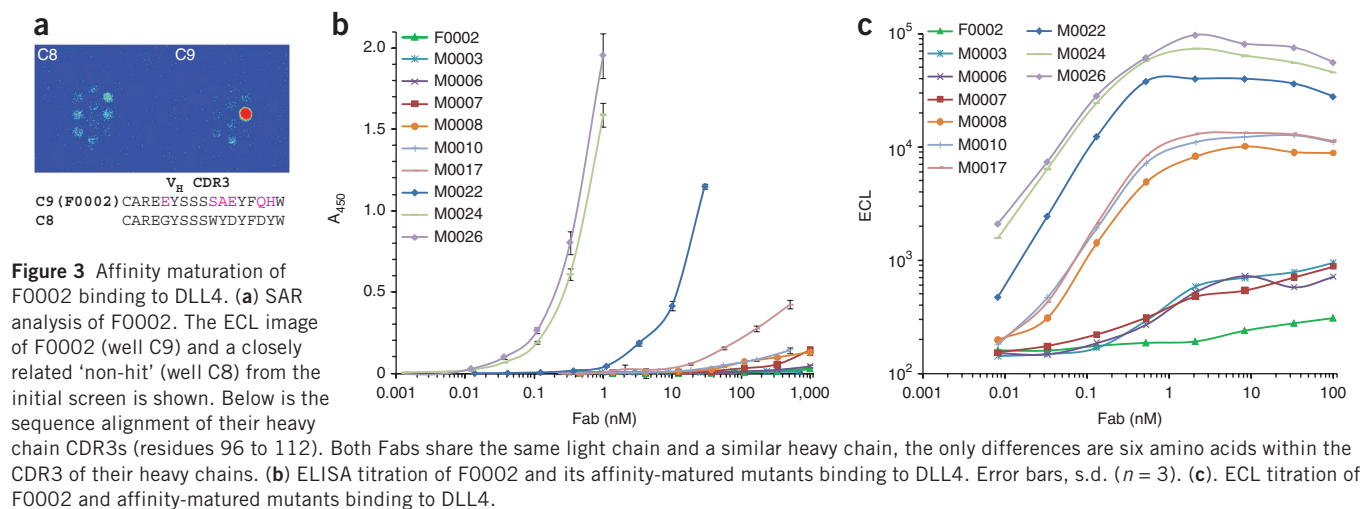
We also applied similar SAR-guided, affinity-maturation principles to the germline hit F0001. By combining five mutations (G24T, S28R, G35V, G100K and G104T) in the heavy chain and three mutations

antibody were both present (without any washing steps). ECL signal was produced only upon Fab binding to a target antigen. We defined a hit as being fourfold above the electrochemiluminescence signal of the blank spot in the same well, with a specific signal defined by binding to only one or two antigens. **Figure 2b** shows ECL image outputs for several specific hits with signal to only one antigen (quantified ECL data are in **Supplementary Table 1**). Polyspecific binding Fabs could also be identified (**Supplementary Fig. 3**). A follow-up titration (**Fig. 2c** and **Supplementary Table 2**) verified the hits and provided information on relative binding strength.

Screening 10,024 Fabs³⁸ using this method revealed 85 hits against seven of nine antigens. The number of specific hits ranged from four against P-cadherin, to 30 against the erythropoietin receptor (EPOR) (**Supplementary Table 3**). We measured the affinity (K_d) of seven of these hits and found them to be generally in the micromolar range as measured by surface plasmon resonance (SPR). The values from highest affinity to lowest were 500 nM (P-cadherin), 730 nM (DLL4), 4.8 μ M (DLL4), 5.6 μ M (EPOR), 8.3 μ M (EPOR), 38 μ M (DLL4) and 65 μ M (ERBB2). These data are consistent with our expectation of identifying low-affinity germline Fabs. Sequence information for the hits revealed that some hits against the same target were derived from similar V, D or J sequences. In other circumstances, disparate V, D or J sequences were identified in Fab hits against the same target. For example, among the hits identified for EPOR binding, we observed four V_H sequences (V_{H1-46} , V_{H4-28} , V_{H4-31} and V_{H3-23}) and ten V_L sequences (B3, L2, A17, A27, L12, O1, O12, L6, V2-6 and V2-17).

Identification of hits against DLL4

One of our target antigens was human delta-like ligand 4 (DLL4), an important ligand in the NOTCH signaling pathway³⁹. DLL4 is a transmembrane protein that is restricted largely to the endothelium of developing vessels and to a small number of additional tissues, such as the thymus, retina, brain, neural tube and hematopoietic cells^{39,40}. Upregulation of DLL4 expression in endothelial vessels in adult humans has been observed in a variety of solid tumors^{41–43}. Inhibition of the DLL4-NOTCH1 interaction has shown antitumor



(S52G, V91L and S96P) in the light chain, we effectively matured the DLL4 binding affinity from 4,800 nM (F0001) to 3 and 5 nM (M0034 and M0035), a 1,000-fold improvement (**Table 1**). During affinity maturation of both F0001 and F0002, the specificity toward DLL4 was confirmed by our multiplex binding analysis. Thus, SARs and directed mutations in the absence of competitive selection can guide protein affinity maturation in a manner analogous to the optimization of small-molecule hits using medicinal chemistry.

Epitope mapping of the anti-DLL4 antibodies

The divergent sequences of F0001 and F0002 suggest they might bind different epitopes. To test this hypothesis, we designed an ECL-based competition assay. Ru-labeled M0008, derived from F0002, served as the detection reagent. We chose unlabeled M0022 and M0031 (**Table 1**), which have similar affinities to DLL4 ($K_{d,s} = 32.7 \pm 11.6$,

36.2 ± 8.5 nM, respectively), as competitors in an equilibrium competition assay. M0022 and M0008 are affinity matured mutants from F0002, and should bind to the same epitope on DLL4; whereas M0031, derived from F0001, uses completely different heavy and light chain genes. **Figure 4a** shows that at equilibrium M0022 competes efficiently with M0008 for binding on DLL4, with a competition K_d of about 30 nM. In contrast, M0031 shows much weaker competition. At 400 nM M0031, only 33% of Ru-labeled M0008 was competed. This result confirmed that M0022 and M0031 have either minimally overlapping or different epitopes on DLL4.

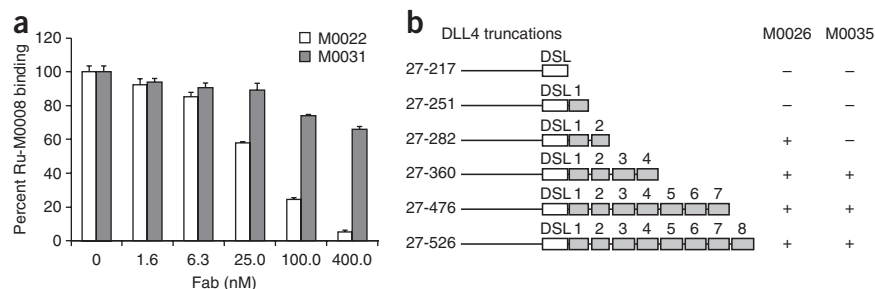
To identify the exact binding location on DLL4, a series of recombinant DLL4 extracellular domain truncations (**Fig. 4b**) with C-terminal myc and His tags were made in Chinese hamster ovary (CHO) cells. The expression and relative sizes of these proteins were confirmed by a western blot probed with anti-myc antibody

Table 1 Affinity-maturation mutations in F0002 and F0001 and their corresponding binding affinities to DLL4

Fab	Residues													k_{on} ($\times 10^5$), $M^{-1}s^{-1}$	k_{off} ($\times 10^{-3}$), s^{-1}	K_d nM	
	H2	H2	H2	H2	H3	H3	H3	H3	L1	L1	L1	L2	L2				
F0002	I	N	S	G	S	S	S	H	S	S	S	S	S	A	1.63	101	730
M0003							F								5.00	190	380
M0006						P									—	—	—
M0007					A										—	—	—
M0008					A	P	F								4.05	49.2	121.5
M0010					A	P	F	F							4.25	30.0	70.6
M0017					A	P	F	F	N	D	H				4.44	68.9	155.2 ^a
M0022				H	A	P	F	F							3.51	10.1	32.7
M0024	V	L	T	H	A	P	F	F	N	D	H				4.30	1.1	2.7
M0026	V	L	T	H	A	P	F	F	N	D	H	L	S		6.84	1.2	1.7
Fab	Residues									k_{on} ($\times 10^5$), $M^{-1}s^{-1}$	k_{off} ($\times 10^{-3}$), s^{-1}	K_d nM					
	H2	H2	H2	H3	H3	L2	L3	L3									
F0001				G	S	G	G	G	S	V	S				—	—	4,800
M0001							K								—	—	—
M0005								T							—	—	—
M0009							K	T							0.6	23	355
M0029					R		K	T							7.4	84.5	114
M0031			T	R	V	K	T	T							20.9	71.7	36.2
M0034			T	R	V	K	T			L	P				110	36	3.3
M0035			T	R	V	K	T		G	L	P				29.6	14.7	5.0

The binding constants are determined by SPR using Bio-Rad ProteOn XPR36. ^aM0017 displayed two-site binding: 89% with K_d of 155.2 nM and 10% with 14 nM.

Figure 4 Epitope mapping the anti-DLL4 Fabs binding sites on the DLL4 extracellular domain. **(a)** ECL-based competition binding assay. Competition between the binding of Ru-labeled M0008 and increased concentrations of either unlabeled M0022 or M0031 to DLL4, quantified by ECL. Error bars, s.d. ($n = 2$). **(b)** Mapping of DLL4 binding region of M0026 and M0035 using an unreduced western blot. A series of DLL4 extracellular domain truncations are depicted in the cartoon diagram, where the DSL domain is highlighted as the open rectangular box; and the EGF-like domains 1 to 8 are highlighted as the shaded rectangular boxes labeled as 1, 2, etc. Anti-DLL4 binding observed in the westerns is marked as +, and no binding is marked as -.



(data not shown). M0026 (derived from F0002) binding involves the EGF2 domain (amino acid (aa) 252 to 286), and M0035 (derived from F0001) binds between EGF2 to EGF4 domains (aa 252 to 360) (Fig. 4b). Both anti-DLL4 Fabs appear to recognize conformational epitopes formed in the presence of disulfide bonds because their binding was abolished when the DLL4 protein was first treated with a reducing reagent.

Functional activity of affinity-matured Fabs

We next sought to evaluate the biological function of the affinity-matured Fabs in the context of the NOTCH1-DLL4 interaction in cell-based assays. First, the binding of M0026 and M0035 to DLL4 overexpressed on the surface of CHO cells (CHO-DLL4) was confirmed using phycoerythrin-labeled secondary antibody (Fig. 5a,b). Biotinylated NOTCH1-Fc binding to CHO-DLL4 was also confirmed using streptavidin-phycoerythrin as the detection reagent (Fig. 5c). Neither Fab showed significant binding to CHO cells without DLL4 overexpression (data not shown). Next, we evaluated whether the DLL4-binding Fabs could compete with NOTCH1-Fc binding to CHO-DLL4. Biotinylated NOTCH1-Fc was incubated with CHO-DLL4 in the presence of different concentrations of unlabeled M0026 or M0035. We confirmed that both anti-DLL4 Fabs could effectively block NOTCH1-Fc binding to CHO-DLL4 (Fig. 5d). Specifically, 2 nM M0026 or 50 nM M0035 competed for more than 80% of NOTCH1 binding activity, with M0026 showing stronger competition than M0035.

We next asked whether M0026 and M0035 could inhibit DLL4-dependent NOTCH1 signaling using a luciferase reporter assay. Human glioma T98G cells⁴⁸, known for the presence of NOTCH1 on their cell surface, were stably transfected with a NOTCH1 reporter plasmid (p6 × CBF) containing six C promoter binding factor-1 (CBF-1) response elements⁴⁹. Firefly luciferase expression was driven through the NOTCH1-DLL4 interaction upon cell contact by addition of CHO-DLL4 cells *in trans* to T98G reporter cells. Incubation of the T98G reporter cells with CHO-DLL4 and an irrelevant Fab resulted in eight- to ninefold increase in NOTCH1 reporter levels compared to those incubated with an irrelevant Fab and CHO cells (Fig. 5e, compare close triangles to open triangles). The NOTCH1 activation remained constant in the presence of increasing concentrations of the non-DLL4 binding Fab F1001 (Fig. 5e). The activation was reduced, however, in the presence of increasing concentrations of anti-DLL4 Fabs M0026 and M0035 (Fig. 5e). The reduction was even more pronounced with an IgG version of M0026 ($IC_{50} \sim 6$ nM) (Fig. 5e), which was almost tenfold more efficient than M0026 Fab. The IgG version of M0035 is also more effective than M0035 Fab, displaying about 30% reduction in NOTCH1 activation at 0.8 nM. However, it did not show complete suppression of NOTCH1 activation at higher concentrations. Clearly, the M0026 IgG is a more potent inhibitor of DLL4-NOTCH1 interaction than M0035 IgG, but only at higher concentrations. Interestingly, M0035 shows partial antagonist activity, being unable to fully inhibit NOTCH1 signaling. Based on these studies, we have demonstrated that target-specific Fabs can be identified from a small germline

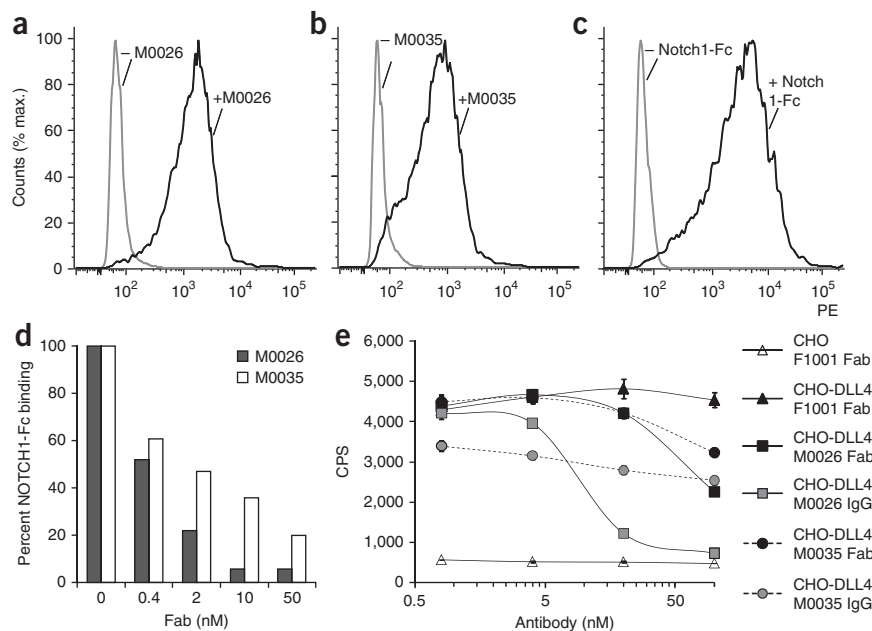


Figure 5 Assays of anti-DLL4 antibodies in cell-based assays. **(a-c)** FACS profiles of M0026, M0035 Fabs and NOTCH1-Fc binding to CHO-DLL4 cells. M0026 and M0035 binding was detected using phycoerythrin-labeled anti-human kappa IgGs (PE- α -kappa), phycoerythrin-labeled anti-human lambda IgG (PE- α -lambda), respectively. Biotinylated NOTCH1-Fc binding to CHO-DLL4 was detected with phycoerythrin-labeled streptavidin (PE-strep). **(d)** Anti-DLL4 Fab competition of NOTCH1-Fc binding to CHO-DLL4 cells using FACS. **(e)** Inhibition of NOTCH1 luciferase reporter by anti-DLL4 antibodies. CHO-DLL4 activation of NOTCH1 reporter in T98G cells is inhibited by both M0026 and M0035 Fabs and IgGs in a dose-dependent manner. A control nonbinding Fab (F1001) has no effect on NOTCH1 reporter levels. Shown is background-subtracted counts per second (CPS) with error bars, s. e. ($n = 4$).

repertoire, these Fabs can be engineered to high-affinity, biologically active molecules and unique activities like partial antagonism can be identified.

DISCUSSION

Although monoclonal antibody discovery dates back 35 years, nearly all FDA approved antibodies are high-affinity antagonists and bind soluble proteins or membrane targets with large extracellular domains. All of these antibodies were discovered using methods relying on biological selection. The ability to discover antibodies with novel mechanisms of action or to derive antibodies against more difficult transmembrane targets, or even unknown targets, could benefit from unique discovery systems that obviate affinity-based competitive selection. Direct screening of collections of individually purified germline antibodies could provide these features.

The feasibility of using arrayed protein libraries for drug discovery has not yet been demonstrated for several reasons. First, the number of library members (paratopes) required to ensure the identification of 'hits' against most targets is unknown. Experience with display libraries suggests this number may be $\sim 10^8$. For example, nonimmune phage libraries of $3 \times 10^7 - 3 \times 10^8$ could select binders in the 0.1–2.5 μM K_d range^{50–52}, whereas libraries of 10^{10} could select binders in the 0.2–70 nM range^{53–55}. The immune system, however, can clearly use a much smaller number of naive antibodies (10^4) as the pool capable of responding to any antigen. Second, the technology to produce large numbers of purified proteins at high yield in parallel was not available until the advent of several high-throughput protein crystallography programs (structural genomics projects)^{56,57}. Third, the cost of DNA synthesis was prohibitive to constructing such large libraries on a one-by-one basis until recently.

The biochemical properties inherent in screening the spatially addressed library reported here are substantially different than in display-based discovery. First, screening is done at high Fab concentrations. Thus, binding can be driven by greater Fab/target ratios in a spatially addressed format using soluble proteins. In contrast, individual members of a display-based library may only be present at attomolar concentrations or less, which could decrease the likelihood of any given binder from interacting with the target, and necessitate larger display-based library sizes. Second, our screening is accomplished in a homogenous binding situation, in the absence of washing events that can remove binders with fast off-rates and further decrease the concentration of the molecule being screened. Third, our discovery method relies on screening as opposed to selection, thus both low- and high-affinity leads can be identified because they are analyzed in separate wells. In this regard, the sensitivity of detection allows identification of specific hits in the 0.5–50 μM range. The ability to detect very low affinity leads may allow for substantially smaller library sizes to be screened successfully. Fourth, germline antibodies are known to adopt many different conformations, with the ability of a single antibody to bind unrelated antigens due to alternative conformations of flexible CDR loops^{21–26,58}. Thus, the actual structural diversity in a library of 10^4 germline antibodies might be significantly higher. In this regard, it has been suggested that the amino acid sequences of V_H CDR3 (encoded largely by the D region) in germline antibodies has been optimized by evolution for maximum flexibility²¹. A small but optimized library for these characteristics might be more efficient at producing binders than other much larger libraries that were generated using *in vivo* affinity-matured V-regions. Lastly, there has been a considerable discrepancy between the apparent library sizes required for discovery of binding proteins using display-based approaches and the actual repertoire size in a mammal that responds to antigen¹⁸.

Because of the selection requirement, and associated washing steps, larger library sizes may be required for display-based discovery. In contrast, we have made a spatially addressed library of a size that is in line with the actual diversity that responds to antigens *in vivo*.

The ability of naive B cells to respond to antigens with a low affinity, followed by subsequent affinity maturation, may allow a small initial repertoire to respond to a multitude of antigens. The ability to achieve robust discovery with such small library sizes in a spatially addressed format could open unique opportunities for biologic discovery of alternative mechanisms of action, novel epitopes, difficult targets, multiplexed screening, multispecific binders or direct screening for functional activity in cell-based assays. Such library sizes (e.g., 10^4) are also well within the limits of most high-throughput discovery efforts using microtiter plates, which make the aforementioned possibilities testable using existing technology. Furthermore, because the sequences are spatially addressed and known a priori, the need to deconvolute the identity of a 'hit' candidate from a single pool of diverse molecules is removed.

A spatially addressed library coupled with the combinatorial nature of V(D)J diversity generation also facilitates identifying SARs that aid in the optimization of preliminary hits, mimicking the approach used in small-molecule medicinal chemistry. Preliminary hits are used to generate an affinity-matured antibody, which contains one or more amino acid alterations that result in an improvement in antigen binding affinity. Currently many of the *in vitro* affinity matured antibodies are produced either by V_H and V_L domain shuffling⁵² or by random mutagenesis of CDR and/or framework residues^{59–64}. However, many of these methods require some type of displayed selection because of the vast number of clones to be evaluated. In contrast, we adopted a more rational and targeted mutagenesis approach, using much smaller libraries guided by SARs and alanine-scanning mutagenesis to identify regions and residues that modulated affinity. Our antibody discovery method can give true SARs because active hits can be compared with related, but inactive, Fabs present in the library. Only residues with alanine substitutions that were in contact-making CDRs but that did not affect binding (or even conferred improvement) were further mutated individually to other amino acids. We also avoided simultaneous mutations to circumvent exponential expansion of the library size. For a given CDR, once the best substitution was identified in each of the mutated positions, they were combined in the new Fab and were found to provide additive improvement in binding affinity. The increase in affinity, measured as a decrease in K_d , can be achieved through either an increase in association rate (k_{on}), a reduction in dissociation rate (k_{off}), or both. Through the homogenous ECL-screening platform, we were able to capture improvements in both directions.

We used *de novo* DNA synthesis and automated parallel fermentation and purification to generate our arrayed library. The time and costs of library construction are greater than generation of 'one pot' libraries used in display approaches. Compared with small-molecule libraries, however, the standardized production and purification techniques in antibody generation allow the costs to be several orders of magnitude below typical combinatorial chemistry collections. These costs and production efficiencies can likely be further optimized through strain development, alternative hosts, enhanced parallel fermentation conditions, optimized automation engineering or even alternative scaffolds to antibodies^{65,66}, as well as the continued decrease in the cost of DNA synthesis. Smaller scale production in 96-well blocks could be accomplished routinely in most molecular biology laboratories using standard strains and micropurification techniques. Furthermore, the discovery efficiency could potentially

be improved by increasing the library size, presenting the antigen in an ordered manner on the solid surface or through miniaturization methods to allow more targets to be screened in smaller volumes³.

Small-molecule drugs have several mechanisms of action, including acting as antagonists, agonists, partial agonists or antagonists and modulators. In contrast, most currently marketed antibodies act as antagonists. The discovery selection mechanisms in hybridoma and display-based systems, which drive affinity and dominant epitope binding, could be reasons for this bias. There is no inherent biochemical reason why antibodies cannot act by alternative mechanisms of action like small-molecule drugs. The M0035 Fab and IgG has partial antagonist activity, and has a qualitatively different activity compared to M0026, which has a similar affinity but binds a different epitope. Discovery of functional activities beyond high-affinity antagonism is a potentially fruitful and largely unexplored area for monoclonal antibody therapeutics. Indeed, unique functional antibodies have been disclosed in a few reports⁶⁷. A limited number of agonist antibodies have been described⁶⁸ and some are now in clinical development⁶⁹. The ability to reproducibly identify such antibodies will require unique discovery systems. The spatially addressed library reported here could enable direct identification of more unique antibodies with novel mechanisms of action beyond antagonism.

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 D. Myszkowski from Biosensor Tools for measuring binding affinities of our Fabs using surface plasmon resonance, R. Lerner for comments on the manuscript, P. Schultz and W. Huse for useful discussions and M. Sandburg and Y. Jiang from Wintherix for their help with the FACS instrument.

AUTHOR CONTRIBUTIONS

H.M. designed and constructed the plasmid vectors for heavy and light chain library; performed Fab library generation; designed and executed affinity maturation, including alanine-scanning mutagenesis, NNK mutagenesis and cassette mutagenesis, of F0001 and F0002; performed a small subset of library screening using ECL; conducted ECL based epitope mapping competition assay. H.M. also designed the cloning strategy for constructing the V_H3-23 library with all possible germline D-J combinations, together with T.M.A.C., J.J.G., C.A.B., O.A.B. and N.P.R. made the V_H3-23 library. J.J.G. performed the automated expression and purification on Piccolo. J.J.G. and V.V.S. designed the software for generating the V(D)J recombinant sequences and selecting the representative sequences for gene synthesis. T.M.A.C. performed the majority of the library screening using ECL and performed ELISA on the affinity matured Fabs with DLL4. C.A.B. performed the Fab binding assays on CHO-DLL4 cells using FACS and executed inhibition assays of NOTCH1-DLL4 interaction using ELISA and FACS. C.A.B. and O.A.B. designed and performed the Luciferase reporter assays on inhibition of NOTCH1-DLL4 interaction. N.P.R. made all DLL4 extracellular domain constructs and executed the epitope mapping with western blots, generated CHO-DLL4 cell line and helped H.M. for Fab library transformation. B.D.S. also supervised the construction of DLL4 extracellular domains. O.A.B. made the NOTCH1 reporter plasmid (p6xCBF), which was modified from an earlier reporter plasmid (p4XCBF) made by B.D.S. O.A.B. also constructed the full length IgG eukaryotic expression vectors and expressed and purified the IgGs. V.V.S. conceptualized the spatially addressed antibody library and oversaw the concept development at Fabrus. V.V.S. and H.M. wrote the manuscript. All authors discussed and commented on 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/>.

- Kohler, G. & Milstein, C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **256**, 495–497 (1975).
- Jin, A. *et al.* A rapid and efficient single-cell manipulation method for screening antigen-specific antibody-secreting cells from human peripheral blood. *Nat. Med.* **15**, 1088–1092 (2009).
- Love, J.C., Ronan, J.L., Grotenbreg, G.M., van der Veen, A.G. & Ploegh, H.L. A microengraving method for rapid selection of single cells producing antigen-specific antibodies. *Nat. Biotechnol.* **24**, 703–707 (2006).
- Hoogenboom, H.R. Selecting and screening recombinant antibody libraries. *Nat. Biotechnol.* **23**, 1105–1116 (2005).
- Huse, W.D. *et al.* Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. *Science* **246**, 1275–1281 (1989).
- McCafferty, J., Griffiths, A.D., Winter, G. & Chiswell, D.J. Phage antibodies: filamentous phage displaying antibody variable domains. *Nature* **348**, 552–554 (1990).
- Hanes, J., Jermutus, L., Weber-Bornhauser, S., Bosshard, H.R. & Pluckthun, A. Ribosome display efficiently selects and evolves high-affinity antibodies in vitro from immune libraries. *Proc. Natl. Acad. Sci. USA* **95**, 14130–14135 (1998).
- Boder, E.T. & Wittrup, K.D. Yeast surface display for screening combinatorial polypeptide libraries. *Nat. Biotechnol.* **15**, 553–557 (1997).
- Beerli, R.R. *et al.* Isolation of human monoclonal antibodies by mammalian cell display. *Proc. Natl. Acad. Sci. USA* **105**, 14336–14341 (2008).
- Ho, M., Nagata, S. & Pastan, I. Isolation of anti-CD22 Fv with high affinity by Fv display on human cells. *Proc. Natl. Acad. Sci. USA* **103**, 9637–9642 (2006).
- Levitani, B. Stochastic modeling and optimization of phage display. *J. Mol. Biol.* **277**, 893–916 (1998).
- Wang, L., Martin, B., Brennen, R., Luttrell, L.M. & Maudsley, S. Allosteric modulators of G protein-coupled receptors: future therapeutics for complex physiological disorders. *J. Pharmacol. Exp. Ther.* **331**, 340–348 (2009).
- Fodor, S.P. *et al.* Light-directed, spatially addressable parallel chemical synthesis. *Science* **251**, 767–773 (1991).
- Ziauddin, J. & Sabatini, D.M. Microarrays of cells expressing defined cDNAs. *Nature* **411**, 107–110 (2001).
- Diller, D.J. The synergy between combinatorial chemistry and high-throughput screening. *Curr. Opin. Drug Discov. Devel.* **11**, 346–355 (2008).
- Polinsky, A. CombiChem and cheminformatics. *Curr. Opin. Drug Discov. Devel.* **2**, 197–203 (1999).
- Kauffman, S.A. *The Origins of Order* (Oxford University Press; 1993).
- Bachmann, M.F., Kundig, T.M., Kalberer, C.P., Hengartner, H. & Zinkernagel, R.M. How many specific B cells are needed to protect against a virus? *J. Immunol.* **152**, 4235–4241 (1994).
- Cohn, M. & Langman, R.E. The protection: the unit of humoral immunity selected by evolution. *Immunol. Rev.* **115**, 11–147 (1990).
- Xu, J.L. & Davis, M.M. Diversity in the CDR3 region of V(H) is sufficient for most antibody specificities. *Immunity* **13**, 37–45 (2000).
- Babor, M. & Kortemme, T. Multi-constraint computational design suggests that native sequences of germline antibody H3 loops are nearly optimal for conformational flexibility. *Proteins* **75**, 846–858 (2009).
- Patten, P.A. *et al.* The immunological evolution of catalysis. *Science* **271**, 1086–1091 (1996).
- Sethi, D.K., Agarwal, A., Manivel, V., Rao, K.V. & Salunke, D.M. Differential epitope positioning within the germline antibody paratope enhances promiscuity in the primary immune response. *Immunity* **24**, 429–438 (2006).
- Thielges, M.C., Zimmermann, J., Yu, W., Oda, M. & Romesberg, F.E. Exploring the energy landscape of antibody-antigen complexes: protein dynamics, flexibility, and molecular recognition. *Biochemistry* **47**, 7237–7247 (2008).
- Yin, J., Beuscher, A.E.t., Andryski, S.E., Stevens, R.C. & Schultz, P.G. Structural plasticity and the evolution of antibody affinity and specificity. *J. Mol. Biol.* **330**, 651–656 (2003).
- Zimmermann, J. *et al.* Antibody evolution constrains conformational heterogeneity by tailoring protein dynamics. *Proc. Natl. Acad. Sci. USA* **103**, 13722–13727 (2006).
- Nguyen, H.P. *et al.* Germline antibody recognition of distinct carbohydrate epitopes. *Nat. Struct. Biol.* **10**, 1019–1025 (2003).
- Thomson, C.A. *et al.* Germline V-genes sculpt the binding site of a family of antibodies neutralizing human cytomegalovirus. *EMBO J.* **27**, 2592–2602 (2008).
- Matsuda, F. *et al.* The complete nucleotide sequence of the human immunoglobulin heavy chain variable region locus. *J. Exp. Med.* **188**, 2151–2162 (1998).
- Smider, V. & Chu, G. The end-joining reaction in V(D)J recombination. *Semin. Immunol.* **9**, 189–197 (1997).
- Gilfilan, S., Benoist, C. & Mathis, D. Mice lacking terminal deoxynucleotidyl transferase: adult mice with a fetal antigen receptor repertoire. *Immunol. Rev.* **148**, 201–219 (1995).
- Holt, L.J., Bussow, K., Walter, G. & Tomlinson, I.M. By-passing selection: direct screening for antibody-antigen interactions using protein arrays. *Nucleic Acids Res.* **28**, E72 (2000).
- Corbett, S.J., Tomlinson, I.M., Sonhammer, E.L., Buck, D. & Winter, G. Sequence of the human immunoglobulin diversity (D) segment locus: a systematic analysis provides no evidence for the use of DIR segments, inverted D segments, “minor” D segments or D-D recombination. *J. Mol. Biol.* **270**, 587–597 (1997).
- Brezinschek, H.P. *et al.* Analysis of the human VH gene repertoire. Differential effects of selection and somatic hypermutation on human peripheral CD5(+)/IgM+ and CD5(-)/IgM+ B cells. *J. Clin. Invest.* **99**, 2488–2501 (1997).

35. Kim, D. *et al.* Directed evolution and identification of control regions of ColE1 plasmid replication origins using only nucleotide deletions. *J. Mol. Biol.* **351**, 763–775 (2005).
36. Leonard, B., Sharma, V. & Smider, V. Co-expression of antibody fab heavy and light chain genes from separate evolved compatible replicons in *E. coli*. *J. Immunol. Methods* **317**, 56–63 (2006).
37. Wollerton, M.C., Wales, R., Bullock, J.A., Hudson, I.R. & Beggs, M. Automation and optimization of protein expression and purification on a novel robotic platform. *JALA* **11**, 291–303 (2006).
38. Smider, V. *et al.* Combinatorial antibody libraries and uses thereof. PCT/US2009/063299 (2009).
39. Shutter, J.R. *et al.* Dll4, a novel Notch ligand expressed in arterial endothelium. *Genes Dev.* **14**, 1313–1318 (2000).
40. Yoneya, T. *et al.* Molecular cloning of delta-4, a new mouse and human Notch ligand. *J. Biochem.* **129**, 27–34 (2001).
41. Li, J.L. *et al.* Delta-like 4 Notch ligand regulates tumor angiogenesis, improves tumor vascular function, and promotes tumor growth in vivo. *Cancer Res.* **67**, 11244–11253 (2007).
42. Mailhos, C. *et al.* Delta4, an endothelial specific notch ligand expressed at sites of physiological and tumor angiogenesis. *Differentiation* **69**, 135–144 (2001).
43. Patel, N.S. *et al.* Up-regulation of endothelial delta-like 4 expression correlates with vessel maturation in bladder cancer. *Clin. Cancer Res.* **12**, 4836–4844 (2006).
44. Hoey, T. *et al.* DLL4 blockade inhibits tumor growth and reduces tumor-initiating cell frequency. *Cell Stem Cell* **5**, 168–177 (2009).
45. Patel, N.S. *et al.* Up-regulation of delta-like 4 ligand in human tumor vasculature and the role of basal expression in endothelial cell function. *Cancer Res.* **65**, 8690–8697 (2005).
46. Segarra, M. *et al.* Dll4 activation of Notch signaling reduces tumor vascularity and inhibits tumor growth. *Blood* **112**, 1904–1911 (2008).
47. Yan, M. *et al.* Chronic DLL4 blockade induces vascular neoplasms. *Nature* **463**, E6–E7 (2010).
48. Nefedova, Y., Cheng, P., Alsina, M., Dalton, W.S. & Gabrilovich, D.I. Involvement of Notch-1 signaling in bone marrow stroma-mediated de novo drug resistance of myeloma and other malignant lymphoid cell lines. *Blood* **103**, 3503–3510 (2004).
49. Purow, B.W. *et al.* Expression of Notch-1 and its ligands, Delta-like-1 and Jagged-1, is critical for glioma cell survival and proliferation. *Cancer Res.* **65**, 2353–2363 (2005).
50. de Kruif, J., Boel, E. & Logtenberg, T. Selection and application of human single chain Fv antibody fragments from a semi-synthetic phage antibody display library with designed CDR3 regions. *J. Mol. Biol.* **248**, 97–105 (1995).
51. Mao, S. *et al.* Phage-display library selection of high-affinity human single-chain antibodies to tumor-associated carbohydrate antigens sialyl Lewisx and Lewisx. *Proc. Natl. Acad. Sci. USA* **96**, 6953–6958 (1999).
52. Marks, J.D. *et al.* By-passing immunization: building high affinity human antibodies by chain shuffling. *Bio/Technology* **10**, 779–783 (1992).
53. Griffiths, A.D. *et al.* Isolation of high affinity human antibodies directly from large synthetic repertoires. *EMBO J.* **13**, 3245–3260 (1994).
54. Sheets, M.D. *et al.* Efficient construction of a large nonimmune phage antibody library: the production of high-affinity human single-chain antibodies to protein antigens. *Proc. Natl. Acad. Sci. USA* **95**, 6157–6162 (1998).
55. Vaughan, T.J. *et al.* Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library. *Nat. Biotechnol.* **14**, 309–314 (1996).
56. Chandonia, J.M. & Brenner, S.E. The impact of structural genomics: expectations and outcomes. *Science* **311**, 347–351 (2006).
57. Joachimiak, A. High-throughput crystallography for structural genomics. *Curr. Opin. Struct. Biol.* **19**, 573–584 (2009).
58. Brooks, C.L. *et al.* Exploration of specificity in germline monoclonal antibody recognition of a range of natural and synthetic epitopes. *J. Mol. Biol.* **377**, 450–468 (2008).
59. Barbas, C.F. III *et al.* In vitro evolution of a neutralizing human antibody to human immunodeficiency virus type 1 to enhance affinity and broaden strain cross-reactivity. *Proc. Natl. Acad. Sci. USA* **91**, 3809–3813 (1994).
60. Cumbers, S.J. *et al.* Generation and iterative affinity maturation of antibodies in vitro using hypermutating B-cell lines. *Nat. Biotechnol.* **20**, 1129–1134 (2002).
61. Hawkins, R.E., Russell, S.J. & Winter, G. Selection of phage antibodies by binding affinity. Mimicking affinity maturation. *J. Mol. Biol.* **226**, 889–896 (1992).
62. Jackson, J.R., Sathe, G., Rosenberg, M. & Sweet, R. In vitro antibody maturation. Improvement of a high affinity, neutralizing antibody against IL-1 beta. *J. Immunol.* **154**, 3310–3319 (1995).
63. Wu, H. *et al.* Stepwise in vitro affinity maturation of Vitaxin, an alpha beta3-specific humanized mAb. *Proc. Natl. Acad. Sci. USA* **95**, 6037–6042 (1998).
64. McCall, A.M. *et al.* Isolation and characterization of an anti-CD16 single-chain Fv fragment and construction of an anti-HER2/neu/anti-CD16 bispecific scFv that triggers CD16-dependent tumor cytolysis. *Mol. Immunol.* **36**, 433–446 (1999).
65. Binz, H.K., Amstutz, P. & Pluckthun, A. Engineering novel binding proteins from nonimmunoglobulin domains. *Nat. Biotechnol.* **23**, 1257–1268 (2005).
66. Skerra, A. Alternative non-antibody scaffolds for molecular recognition. *Curr. Opin. Biotechnol.* **18**, 295–304 (2007).
67. Miller, D.J. & Rodriguez, M. A monoclonal autoantibody that promotes central nervous system remyelination in a model of multiple sclerosis is a natural autoantibody encoded by germline immunoglobulin genes. *J. Immunol.* **154**, 2460–2469 (1995).
68. Liu, Z. *et al.* A potent erythropoietin-mimicking human antibody interacts through a novel binding site. *Blood* **110**, 2408–2413 (2007).
69. Wang, J. *et al.* Characterization of a novel anti-DR5 monoclonal antibody WD1 with the potential to induce tumor cell apoptosis. *Cell. Mol. Immunol.* **5**, 55–60 (2008).

ONLINE METHODS

Library construction. Plasmids A, C and E were constructed to produce the recombinant Fab library. Plasmid A was used for expressing the heavy chains, plasmid C for κ -light chains, and plasmid E for λ light chains. The ColE1-derived replication origins for plasmids A and C (or plasmids A and E) are compatible to each other and replicate at similar copy numbers when co-cultured in the same *Escherichia coli* cell. The plasmids carry a STII leader sequence for Fab secretion and an Ara promoter for inducible gene expression. Plasmid A encodes a C_H region for production of a Fab heavy chain and includes Flag and His tags for protein purification. Plasmid C encodes a C_κ region for production of a κ -Fab light chain while plasmid E encodes a C_λ region for production of a λ -Fab light chain. *In silico* V(D)J recombination of germline antibody gene sequences was performed using proprietary software developed at Fabrus. DNAs encoding the recombined V_H or V_L regions were synthesized by GenScript. The V_H sequences were cloned between the NheI and NcoI sites of plasmid A. The V_κ sequences were cloned between the NcoI and BsiWI of plasmid C, and the V_λ sequence was cloned between the NcoI and AvrII sites of plasmid E using standard molecular biology techniques.

A Fab is produced in *E. coli* upon co-transformation and induction of a plasmid encoding a heavy chain and a plasmid encoding a light chain. Specifically, 1 ng heavy and light chain plasmids were combined in a PCR tube or a PCR 96-tube plate and were mixed well with 20 μ l ice cold LMG194 (ATCC)-competent cells. After heat-shock transformation and recovery at 37 °C, the co-transformed cells were selected by growing in 1 ml Luria-Bertani broth containing 0.4% (wt/vol) glucose (Sigma Aldrich), 50 μ g/ml kanamycin (Sigma Aldrich) and 34 μ g/ml chloramphenicol (Sigma Aldrich) at 32 °C with vigorous shaking for 20 h. To create a Fab genetic library, different heavy chains in 96-well format are co-transformed with multiple light chains in parallel. Sequences of all Fabs in the library are available in ref. 38.

Automated expression and purification. Overnight co-transformation cultures in 96-deep well format were used as inocula for Piccolo instrument (The Automation Partnership) expression and purification of the Fabs³⁷. All Piccolo labware was barcoded for tracking through the process. Five liters of sterile Terrific Broth (TB) media was prepared with kanamycin (50 μ g/ml), chloramphenicol (35 μ g/ml), glucose (2% wt/vol) and antifoam 204 (0.015% vol/vol) (Sigma). Arabinose inducer (0.2% wt/vol) (Sigma) and the prepared media were attached to sterile dispensing pumps. Culture vessel blocks (CVBs) were filled (9.5 ml of media) and inoculated (100 μ l per well). After all CVBs were inoculated they were transferred one-by-one to an incubator and growth was monitored by an OD₆₀₀ reading. Upon completion, a CVB was removed from the incubator and 2.5 ml lysis buffer (50 mM sodium phosphate pH 8.0, 300 mM NaCl, 3.25% wt/vol N-octyl- β -D thioglucoside (Alexis Biochemicals) and 1:500 dilution of Lysonase (EMD)) was added to each well. Lysis proceeded for 30 min and then the CVB was centrifuged for 30 min at 5,000g to pellet the cellular debris. The supernatant was transferred to a filter plate containing 1 ml His-Bind resin (EMD) per well. The resin was then washed 2 \times with 6 ml of 50 mM sodium phosphate pH 8.0, 300 mM NaCl, 30 mM imidazole and eluted 4 \times with 500 μ l of 50 mM sodium phosphate pH 8.0, 300 mM NaCl, 500 mM imidazole.

Fabs eluted from the His-Bind resin on Piccolo were further purified using a secondary off-line purification step. Depending on the light chain classes, two different affinity resins were applied: Fabs with κ -light chain were further purified on Protein G column (GE Healthcare), and those with λ -light chain were purified on CaptureSelect Fab Lambda affinity column (BAC). Approximately 1.8 ml of the His-Bind elution per Fab sample was purified on either a 1 ml Hi-Trap Protein G column or a 0.5 ml CaptureSelect Fab Lambda affinity column at 4 °C using the Akta purifier (GE Healthcare) and A-905 autosampler (GE Healthcare) according to the manufacturer's protocol. After four 450 μ l injections, the column was washed with two column volumes of 50 mM sodium phosphate pH 7.2, 150 mM NaCl. The Fab was eluted with six column volumes of 100 mM glycine (EMD), pH 2.8. Approximately 0.8 ml of elution peak fractions were collected in a deep well, 96-well plate block, and immediately neutralized with 100 μ l saturated dibasic sodium phosphate pH 9.0. The Fab concentration was determined from A₂₈₀ absorbance on a plate reader (Molecular Devices).

Fab library screening. Nine recombinant human antigens (carrier free) were purchased from R&D systems: ERBB2/Fc, EGFR/Fc, HGFR/Fc, NOTCH1/Fc,

CD44/Fc, IGF1R, P-cadherin/Fc, EPOR/Fc and DLL4. These antigens were spotted onto 96-well Multi-Spot 10 Highbind plates by Meso Scale Discovery (MSD). Spots 1 to 9 contain 50 nl of 60 μ g/ml of antigen each, and spot 10 is left blank. Goat anti-human κ -light chain polyclonal antibody (Sigma Aldrich) and Goat anti-human λ light chain polyclonal antibody (Sigma Aldrich) were conjugated with Ru tri-bispyridine-(4-methylsulfone) TAG (MSD) according to the manufacturer's suggested procedures.

The antigen-coated, 96-well Multi-Spot 10 Highbind plate was first blocked with 150 μ l of 3% bovine serum albumin (BSA) in Tris-buffered saline containing 0.1% Tween-20 (TBST) per well at 20 °C for 30 min. After washing, 12.5 μ l 1% BSA TBST and 12.5 μ l of Fab from the purified library were added per well and incubated at 20 °C with shaking for 1 h. Ru-labeled detection antibody (1 μ g/ml) was added in 25 μ l and incubated at 20 °C with shaking for 1 h. Finally 15 μ l of Buffer P with surfactant (MSD) was added to the complex mixture and the plate was read on a Sector Imager 2400 (MSD). The ECL signals of the antigen and the blank in each well were compared and a signal/blank ratio of 4:1 or greater was considered a potential hit. To confirm a potential hit from the initial ECL screening, a secondary titration with sequential dilution was carried out.

Epitope mapping. Single-spot, 96-well standard plates (MSD) were coated overnight with 5 μ l per well of 10 μ g/ml DLL4 (R&D systems) in PBS with 0.03% Triton X-100. Control wells were left uncoated as blank. The next day, a 150 μ l aliquot of 3% BSA in TBST was added to each well and allowed to incubate for 60 min at 20 °C to block the plate. After washing twice with 150 μ l TBST and tap drying, 50 μ l aliquots of 100 nM Ru-labeled M0008 plus a serial dilution of either M0022 or M0031, in duplicates, were added. After incubating the plate at 20 °C with shaking for 1 h, ECL was measured using a Sector Imager 2400. After subtracting the signals from the blank wells without DLL4, percent Ru-M0008 binding was calculated by dividing the ECL signal to the averaged signal with only Ru-labeled M0008.

Alanine-scanning mutagenesis and NNK mutagenesis. Oligos used for alanine-scanning mutagenesis were synthesized by Integrated DNA Technologies (IDT). The mutagenesis was carried out using standard overlapping PCR procedures. NNK mutagenesis by overlapping PCR was carried out similarly to the alanine scanning mutagenesis, except the target codon is replaced with NNK in the forward primer, and MNN in the reverse primer. After transformation of the ligation product in DH5 α and plating, individual colonies were picked into 96-well blocks containing 1.5 ml of Terrific Broth (EMD) supplemented with 50 μ g/ml kanamycin and 0.4% glucose per well, and grown at 37 °C overnight for isolating the plasmid of each mutation. DNA sequencing was used to identify the exact substitution of G100 in each well.

Cassette mutagenesis. We designed a straightforward cloning strategy to carry out cassette mutagenesis using Type II restriction enzymes such as BsaI. First, internal BsaI sites (if any) encoded by the plasmid were removed using the Quick Change Mutagenesis Kit (Stratagene). Depending on the CDR to be mutated, a vector was made to incorporate a BsaI site at the beginning and at the end of the CDR using overlapping PCR. The vector was then digested with BsaI and gel purified. A set of short forward and reverse primers carrying the desired mutation in the CDR were synthesized. When annealed, they formed compatible ends to the digested vector. The primer sets were mixed in TE ((tris[hydroxymethyl]-aminomethane), ethylenediaminetetraacetic acid) at 1 μ M, heated at 95 °C for 2 min and slowly cooled down to 24 °C. One microliter of the annealed primers was ligated with 2 ng of the BsaI digested vector and transformed into DH5 α cells. As long as the mutations were within the same CDR, the same vector was used for incorporating any mutations by ligation with short primer sets.

Enzyme-linked immunosorbent assay (ELISA) titration of anti-DLL4 Fabs. Nunc Maxisorp 384-well plates were coated with 10 μ l per well of 0.5 μ g/ml recombinant human DLL4 extracellular domain overnight at 4 °C. After blocking the plate with 1% BSA in TBST, a 20 μ l aliquot of each serial dilution was added, in triplicate, to each well and the plate was incubated for 1 h at 24 °C followed by washing 2 times with 100 μ l TBST. Depending on the light chain (either κ or λ), 20 μ l of the corresponding goat anti-human kappa horseradish

peroxidase (HRP) conjugated polyclonal antibody (Sigma-Aldrich) or goat anti-human lambda HRP conjugated polyclonal antibody (Sigma-Aldrich) diluted 1:1,000 in 1% BSA TBST was added to each well and the plate was incubated for 1 h at 24 °C followed by washing 4× with 100 µl TBST. Finally, 25 µl TMB one-component reagent (BioFfx) was added and allowed to develop for 1 to 10 min at 24 °C. The reaction was immediately stopped by the addition of 25 µl 0.5 M H₂SO₄ and the absorbance at 450 nm was measured on a plate reader (Molecular Devices).

K_d measurement by SPR. Binding studies were outsourced to Biosensor Tools. They were run on a ProteOn system using a GLM sensor chip in 10 mM HEPES (pH 7.4), 150 mM NaCl, 0.01% tween-20 and 0.1 mg/ml BSA at 25 °C. Each antigen was immobilized using amine coupling at three different surface densities. The Fabs were each tested at 1 µM as the highest concentration in a threefold dilution series. The response data from each surface were globally fit to determine the binding constants presented on each plot and averaged for the final report.

IgG cloning, expression and purification. Sequences encoding heavy and light chains were cloned separately into the pFUSE family of vectors (InvivoGen). To produce IgG, the heavy and light chain plasmids were co-transfected into 293FS cells (Invitrogen) using 293fectin (Invitrogen) per manufacturer's instructions. Cells growing in serum-free 293Freestyle medium (Invitrogen) were transfected at 1 × 10⁶ cells/ml in a 50 ml spinner flask. Cell culture media was harvested 3 and 6 d after transfection and pooled together for purification on Protein-G Sepharose (GE Healthcare). IgG elution fractions were pooled and dialyzed into PBS.

DLL4 extracellular domain and DLL4 truncations preparation. Human DLL4 cDNA (Open Biosystems) was used as the template for PCR amplification of the full extracellular domain of DLL4 (aa 1–524), and several extracellular domain truncations: pre-DSL (aa 1–172), DSL (aa 1–217), EGF1 (aa 1–251), EGF2 (aa 1–286), EGF4 (aa 1–360) and EGF7 (aa 1–476). A myc tag and a 6-His tag were appended to the C termini of these constructs. The PCR products were cloned between the NheI and NotI sites of pcDNA5/FRT vector (Invitrogen) and transfected into the Flp-In CHO Cell Line (Invitrogen) according to the Flp-In System protocol. Transfected cells were selected with 400 µg/ml hygromycin after 2 d. After colony picking, cell lines were maintained at 37 °C with hygromycin selection. Expression media samples were collected after 7 d. One ml aliquots were batch bound with 50 µl of Talon Resin (Clontech) for 30 min. Samples were washed, boiled with loading dye ± DTT, and resolved on a 4–12% Bis-Tris Gel (Invitrogen). Protein bands were transferred to PVDF membrane and probed with either 0.5 µg/ml anti-c-myc mAb (Genscript) to confirm protein secretion, or 20 nM M0026 or M0035 for epitope mapping. For detection goat anti-mouse HRP, anti-human kappa HRP, or anti-human lambda HRP (1,000-fold dilution) were used, respectively.

CHO-DLL4 binding and inhibition of DLL4-NOTCH1 interaction using flow cytometry. Full-length human DLL4 gene was digested out from the OpenBiosystems vector and ligated directly into pcDNA5/FRT between NheI and NotI sites. Transfection and colony selection were as described above.

CHO cells stably expressing full-length DLL4 or control CHO cells were detached from plates using Accutase (eBiosciences). Cells were washed in 2% BSA/PBS and incubated with 10–50 nM Fab in 2% BSA/PBS for 30 min. Phycoerythrin-labeled anti-human kappa (Invitrogen) for M0026 or phycoerythrin-labeled anti-human lambda (Invitrogen) for M0035 at a dilution of 1:30 were added to cells and incubated for 10 min. Cells were then washed in 2% BSA/PBS and analyzed by flow cytometry on a FACSaria (Becton Dickinson).

To test Fab inhibition of the NOTCH1-DLL4 interaction, cells were treated with 250 nM biotinylated NOTCH1-Fc and 0 to 50 nM Fab (either M0026 or M0035) for 30 min in 2% BSA/PBS. Phycoerythrin-labeled streptavidin (Pierce-Thermo Scientific) was then added to a final dilution of 1:5 and then incubated for 10 min. Cells were then washed in 2% BSA/PBS and analyzed by flow cytometry on a FACSaria.

Inhibition of NOTCH1 signal transduction. A NOTCH1 reporter construct (p6 × CBF) was made by inserting six CBF NOTCH-response elements (bold) into pGL4.26 (Promega): GGTACCTGAGCTCGCTAGCGATCTGGTGTAACACGCCGTGGGAAAAAATTTATGGATCTGGTGTAACACGCCGTGGGAAAAAATTTATGGAGCTCGCTAGCGATCTGGTGTAACACGCCGTGGGAAAAAATTTATGGATCTGGTGTAACACGCCGTGGGAAAAAATTTATGGATCTGGTGTAACACGCCGTGGGAAAAAATTTATGGAGCTT. The reporter plasmid was transfected into T98G glioma cells (ATCC), which express endogenous NOTCH1. Stable integrants were selected with 200 µg/ml Hygromycin B (Invitrogen), yielding a T98G NOTCH1 reporter cell line that induced Firefly luciferase expression upon NOTCH1 activation. CHO cells expressing DLL4 or control CHO cells were propagated in F12 media (Invitrogen) supplemented with 10% FBS and P/S/G. Separately, T98G NOTCH1 reporter cells (2 × 10⁵ cells/well) in EMEM with 10% FBS and P/S/G were plated onto 96-well tissue culture plates. The next day, media was removed from the T98G cells and replaced with 100 µl CHO-DLL4 or control CHO cells (1 × 10⁵ cells/well) in serum free F12 media supplemented with P/S/G. Inhibitory Fabs (M0026 and M0035) and their corresponding IgGs, and control Fabs (F1001) were added at 100, 20, 4 and 0.8 nM. After 24 h, luciferase-reporter expression was measured with Bright-Glo luciferase assay reagent (Promega). Luminescence was read using a Wallac Victor II model 1420 plate reader. Each condition was performed in quadruplicate. CHO cells (not expressing DLL4) were never seen to activate NOTCH1 reporter alone or in combination with any antibody (Fig. 5e and data not shown).

Antibody recycling by engineered pH-dependent antigen binding improves the duration of antigen neutralization

Tomoyuki Igawa, Shinya Ishii, Tatsuhiko Tachibana, Atsuhiko Maeda, Yoshinobu Higuchi, Shin Shimaoka, Chifumi Moriyama, Tomoyuki Watanabe, Ryoko Takubo, Yoshiaki Doi, Tetsuya Wakabayashi, Akira Hayasaka, Shoujiro Kadono, Takuya Miyazaki, Kenta Haraya, Yasuo Sekimori, Tetsuo Kojima, Yoshiaki Nabuchi, Yoshinori Aso, Yoshiki Kawabe & Kunihiro Hattori

For many antibodies, each antigen-binding site binds to only one antigen molecule during the antibody's lifetime in plasma. To increase the number of cycles of antigen binding and lysosomal degradation, we engineered tocilizumab (Actemra)¹, an antibody against the IL-6 receptor (IL-6R), to rapidly dissociate from IL-6R within the acidic environment of the endosome (pH 6.0) while maintaining its binding affinity to IL-6R in plasma (pH 7.4). Studies using normal mice and mice expressing human IL-6R² suggested that this pH-dependent IL-6R dissociation within the acidic environment of the endosome resulted in lysosomal degradation of the previously bound IL-6R while releasing the free antibody back to the plasma to bind another IL-6R molecule. In cynomolgus monkeys, an antibody with pH-dependent antigen binding, but not an affinity-matured variant, significantly improved the pharmacokinetics and duration of C-reactive protein inhibition. Engineering pH dependency into the interactions of therapeutic antibodies with their targets may enable them to be delivered less frequently or at lower doses.

Therapeutic antibodies now offer important treatment options for many diseases^{3,4}. Although most antibodies effectively neutralize their cognate antigens, relatively high doses and/or frequent injections are required to achieve therapeutic efficacy when targeting antigens, such as IL-6R⁵, EGFR⁶, IgE⁷ CD4 (ref. 8) and CD40 (ref. 9), with high rates of synthesis. Antibodies targeting these antigens show significant antigen-mediated clearance. For many conventional antibodies, each antigen-binding site binds to only one antigen molecule during the antibody's lifetime in plasma. Because their cognate antigens are usually produced continuously *in vivo*, the minimum number of antigen-binding sites required to completely neutralize the function of the antigen cannot be less than the amount of antigen that is produced between doses. This restricts the dose and dosing frequency of many therapeutic antibodies. For most antibodies, improving the binding affinity¹⁰ and the pharmacokinetics by increasing binding to the neonatal Fc receptor (FcRn)^{11–13} is, to some extent, effective for reducing the therapeutic dose and/or the dosing frequency. Nonetheless, the extent of such

improvements hits a ceiling defined by the limited number of antigen-binding events per antibody molecule and the rate at which the antigen continues to be produced. Further improvement can only be achieved by increasing the number of cycles in which an antibody binds to and releases the antigen for subsequent lysosomal degradation and thereby increasing the amount of antigen that can be bound by a single antigen-binding site. To date, although various technologies for optimizing antibody therapeutics have been reported¹⁴, there is no technology for increasing the number of cycles of antigen binding and lysosomal degradation.

Tocilizumab, a humanized IgG antibody against IL-6R that is approved for treating moderate to severe rheumatoid arthritis¹, exhibits a high rate of antigen-mediated clearance *in vivo* as a result of the high turnover of IL-6R following its delivery⁵. There are both membrane-bound (mIL-6R) and soluble (sIL-6R) forms of IL-6R. Both transduce IL-6 signaling by means of gp130 (ref. 15) and both are targeted by tocilizumab. We engineered unmodified tocilizumab¹⁶ (TCZ) to dissociate from IL-6R within the acidic environment of the endosome by decreasing its binding affinity at pH 5.5–6.0 (ref. 17), without affecting its binding affinity to IL-6R in plasma at pH 7.4. Previous studies have shown that engineering cytokine ligands to dissociate from their receptors rescued the ligands from lysosomal degradation *in vitro*^{18,19}. Therefore, we expected that enabling dissociation of the antibody from IL-6R within the endosome would rescue the antibody from lysosomal degradation, enabling free antibody to be recycled back to the cell surface (for release to the plasma *in vivo*) in order to bind to another IL-6R molecule.

We subjected all residues in the complementarity determining regions and several framework residues important for antigen binding to histidine scanning to identify histidine mutations that confer pH-dependent binding of IL-6R. Histidine, which has a pK_a of approximately pH 6, is often found at the interfaces between proteins that interact with different affinities in plasma and endosomes²⁰. Surface plasmon resonance (SPR) identified several histidine mutations that conferred pH-dependent binding of TCZ to IL-6R. Separate mutation of Tyr27, Ser31 and Trp35 to His in the heavy chain variable region and separate mutation of Asp28, Tyr32, Arg53 to His and His55 to Leu in the light chain variable

Chugai Pharmaceutical Co. Ltd., Fuji-Gotemba Research Laboratories, Shizuoka, Japan. Correspondence should be addressed to T.I. (igawatmy@chugai-pharm.co.jp).

Received 31 August; accepted 23 September; published online 17 October 2010; doi:10.1038/nbt.1691

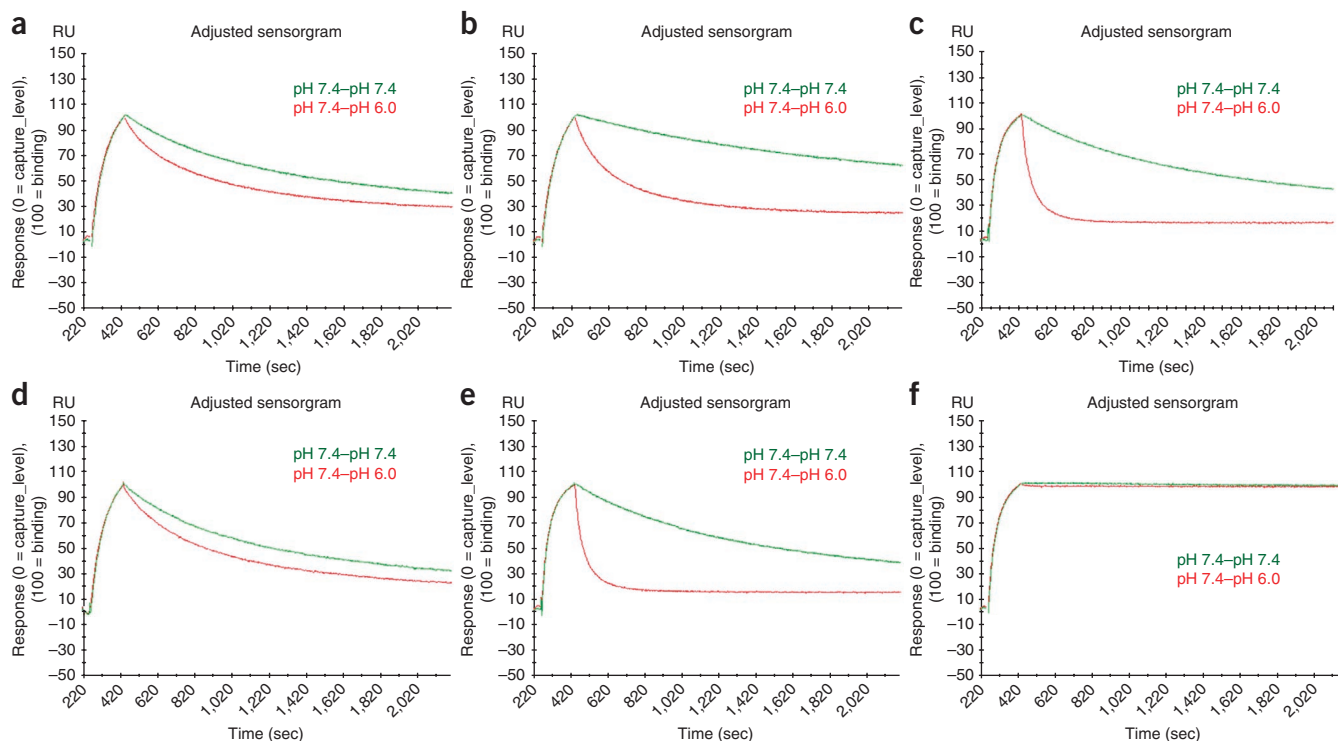


Figure 1 Surface plasmon resonance (SPR) sensorgrams of tocilizumab (TCZ), two variants with pH-dependent binding to hsIL-6R (PH1, PH2), TCZ and PH2 with increased affinity to FcRn (TCZ-FcRn, PH2-FcRn) and an affinity matured variant with increased affinity to FcRn (AM-FcRn). (a–f) SPR measurements of hsIL-6R association (3 min) at pH 7.4 and dissociation (30 min) at pH 7.4 (pH 7.4–pH 7.4; green), and hsIL-6R association at pH 7.4 (3 min) and dissociation at pH 6.0 (30 min) (pH 7.4–pH 6.0; red) are shown for TCZ (a), PH1 (b), PH2 (c), TCZ-FcRn (d), PH2-FcRn (e) and AM-FcRn (f).

region influenced the pH dependency of IL-6R binding (positions numbered in Kabat numbering). A combination of various mutations including mutations at Tyr27, Ser31 and Trp35 in the heavy chain and Asp28, Tyr32 and Arg53 in the light chain generated the TCZ variant PH1, and mutations at Tyr27 and Ser31 in the heavy chain and Tyr32, Arg53 and His55 in the light chain generated PH2. SPR measurements of all TCZ variants used in this study are shown in **Figure 1**. The results indicated that the respective binding affinities of PH1 and PH2 to human soluble IL-6R (hsIL-6R) at pH 7.4 were 2.3- and 2.8-times stronger than that of TCZ. In contrast, their respective binding affinities at pH 6.0 were substantially less than that of TCZ (3.8- and 3.7-times lower). Ratios of binding affinities determined at pH 6.0 and pH 7.4 (K_D s) were ~19 and ~22 for PH1 and PH2, respectively, compared to ~2 for TCZ (**Table 1**). After binding to hsIL-6R at pH 7.4, subsequent rapid dissociation from hsIL-6R at pH 6.0 was observed for PH1 and PH2, but not for TCZ (**Fig. 1a–c**).

As TCZ, PH1 and PH2 do not bind to mouse IL-6R, they were administered intravenously to transgenic mice expressing human IL-6R (hIL-6R)². We measured serum amyloid A (SAA) as a pharmacodynamic marker of IL-6R neutralization²¹ after intraperitoneal injection of human IL-6 (hIL-6). This model exhibits an hIL-6-dependent increase in SAA, which was inhibited by TCZ at plasma TCZ concentrations >10 $\mu\text{g/ml}$ (**Supplementary Fig. 1a–b**). We obtained time profiles for changes in the plasma concentrations of antibody (**Fig. 2a**) and SAA (**Fig. 2b**), as well as the associated pharmacokinetic parameters (**Supplementary Table 1**). The PH1 and PH2 variants of TCZ, which exhibit pH-dependent binding to IL-6R, exhibited a significant improvement in pharmacokinetics and duration of SAA inhibition compared to TCZ. The slight improvement for PH2 compared to PH1 may be due to more rapid dissociation from IL-6R at pH 6.0 (larger k_d at pH 6.0) or to greater pH dependency of binding affinity (larger K_D ratio of pH 6.0/pH 7.4) (**Table 1**). In this study group, the increase in SAA levels could not be inhibited by TCZ at 54 h because

Table 1 Binding kinetics of TCZ, pH-dependent binding variants and affinity matured variant to hsIL-6R

Antibody	pH 7.4			pH 6.0			k_d (pH 6.0) ratio vs. TCZ	K_D ratio pH 6.0/pH 7.4
	k_a ($\text{M}^{-1}\text{s}^{-1}$)	k_d (s^{-1})	K_D (M)	k_a ($\text{M}^{-1}\text{s}^{-1}$)	k_d (s^{-1})	K_D (M)		
TCZ	$2.2\text{E} + 05$	$1.1\text{E} - 03$	$5.1\text{E} - 09$	$2.1\text{E} + 05$	$2.3\text{E} - 03$	$1.1\text{E} - 08$	–	2.1
PH1	$2.6\text{E} + 05$	$5.7\text{E} - 04$	$2.2\text{E} - 09$	$1.0\text{E} + 05$	$4.3\text{E} - 03$	$4.2\text{E} - 08$	1.9	18.9
PH2	$5.3\text{E} + 05$	$1.0\text{E} - 03$	$1.9\text{E} - 09$	$3.5\text{E} + 05$	$1.5\text{E} - 02$	$4.1\text{E} - 08$	6.3	22.1
TCZ-FcRn	$2.8\text{E} + 05$	$1.3\text{E} - 03$	$4.7\text{E} - 09$	$2.1\text{E} + 05$	$2.6\text{E} - 03$	$1.2\text{E} - 08$	1.1	2.6
PH2-FcRn	$5.6\text{E} + 05$	$1.0\text{E} - 03$	$1.8\text{E} - 09$	$3.8\text{E} + 05$	$1.5\text{E} - 02$	$4.0\text{E} - 08$	6.6	22.1
AM-FcRn	$5.7\text{E} + 05$	$2.9\text{E} - 05$	$5.2\text{E} - 11$	$7.8\text{E} + 05$	$1.4\text{E} - 05$	$1.8\text{E} - 11$	0.006	0.4

Association rate (k_a), dissociation rate (k_d) and binding affinity (K_D) of TCZ, PH1, PH2, TCZ-FcRn, PH2-FcRn and AM-FcRn to hsIL-6R at pH 7.4 and pH 6.0, their TCZ to variant k_d at pH 6.0 ratio, and their pH 6.0 to pH 7.4 binding affinity (K_D) ratio are shown.

the plasma TCZ concentration at 54 h was <10 $\mu\text{g/ml}$. A slightly higher SAA concentration in the TCZ group compared to the control (IL-6 + vehicle) group at 78 h most likely resulted from a lower SAA concentration for the control group in this specific study group (normally the concentration of SAA for the control group is ~ 200 $\mu\text{g/ml}$).

As IL-6R exists in membrane-bound and soluble forms¹⁵, we used normal mice for separate evaluation of the effects of pH-dependent dissociation from mIL-6R and sIL-6R. The effect of pH-dependent binding on mIL-6R was evaluated by comparing the pharmacokinetics of TCZ and PH2 in hIL-6R transgenic mice and normal mice. In normal mice, both TCZ and PH2 exhibited longer plasma persistence than in hIL-6R transgenic mice because of the absence of human mIL-6R-mediated clearance (Fig. 2c and Supplementary Table 2). Although the profile of changes in the plasma concentration of PH2 was significantly better than that of TCZ in hIL-6R transgenic mice (20-fold higher plasma concentration at 78 h), the levels were comparable in normal mice (1.2-fold higher plasma concentration on day 14). This demonstrates that the improvement had been mainly due to the reduced clearance mediated by human mIL-6R. In other words, TCZ binds to mIL-6R and is removed from plasma by internalization and subsequent lysosomal degradation (that is, mIL-6R-mediated clearance). Degradation of antibody as a complex with mIL-6R allows a single antigen binding site to bind to only one mIL-6R molecule during its lifetime in plasma (Supplementary Fig. 2a). On the other hand, the improved pharmacokinetics exhibited by PH2 in hIL-6R transgenic mice suggests that mIL-6R-mediated clearance is reduced because the antibody is recycled back to plasma after being internalized by the mIL-6R. After it dissociates from mIL-6R in the acidic endosome, the antibody is recycled back to the plasma and can bind to another mIL-6R molecule. In contrast, the mIL-6R in the endosome is targeted for lysosomal degradation, thereby allowing a single antibody to repeatedly neutralize multiple mIL-6R molecules (Supplementary Fig. 2b).

The effect of pH-dependent binding on sIL-6R was evaluated by administration of hsIL-6R to normal mice either on its own or together with TCZ or PH2. Considering the binding affinity and amount of TCZ and PH2, we predicted that 99.6% and 99.9% of hsIL-6R should be bound to the antibody in the administration solution. A time profile of changes in hsIL-6R plasma concentration and the associated pharmacokinetic parameters are shown in Figure 2d and Supplementary Table 3, respectively. The binding of soluble antigen to antibody is reported to reduce the clearance of antigen as the antibody, with a long half-life, serves as a carrier for the antigen²². As expected, although hsIL-6R alone showed rapid clearance, clearance of hsIL-6R was reduced by 39-fold when hsIL-6R was administered together with TCZ. In contrast, although clearance for TCZ and PH2 were comparable (Supplementary Table 2), the clearance of hsIL-6R in complex with PH2 was enhanced sixfold compared to hsIL-6R in complex with TCZ, resulting in a 40-fold lower concentration of hsIL-6R in plasma on day 4. As the binding affinity of PH2 is 2.8-fold stronger than that of TCZ at pH 7.4, increased clearance of

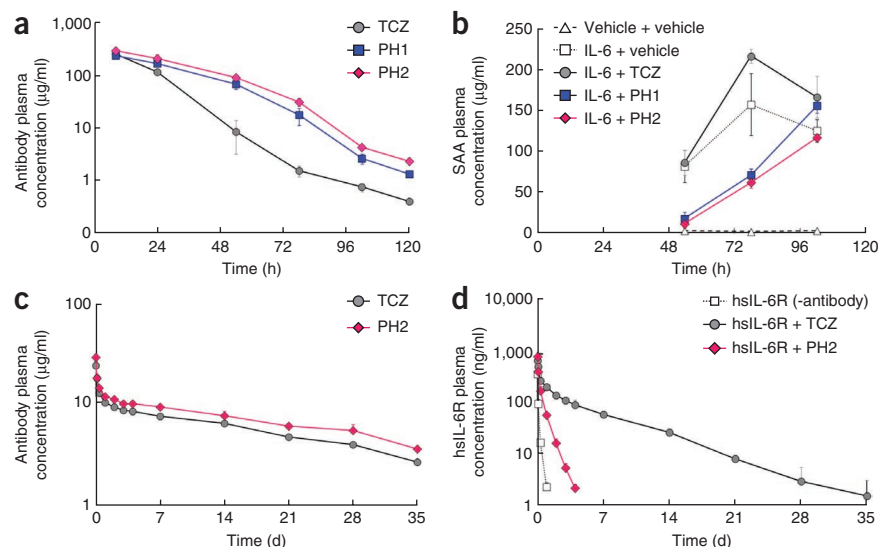
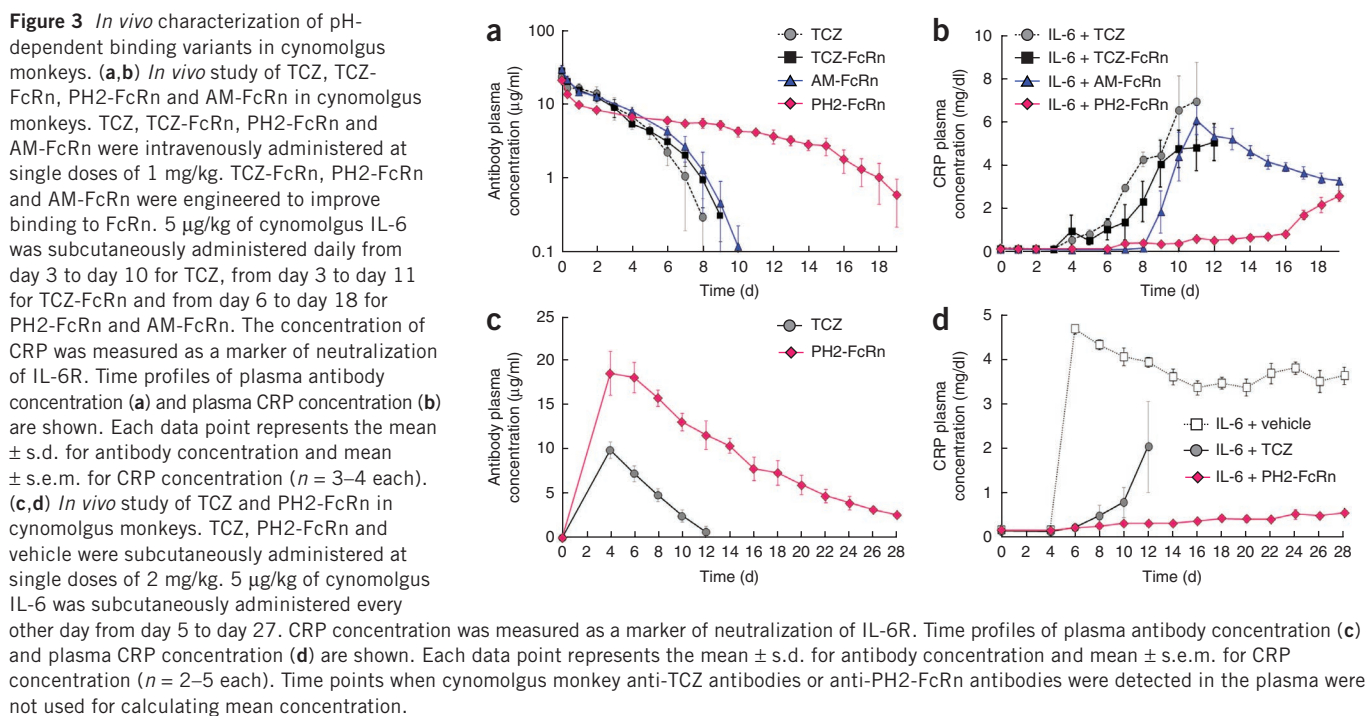


Figure 2 *In vivo* characterization of pH-dependent binding variants in mice. (a,b) *In vivo* study of TCZ, PH1 and PH2 in hIL-6R transgenic mice. TCZ, PH1 and PH2 were administered intravenously at single doses of 25 mg/kg. 4 $\mu\text{g/kg}$ of hIL-6 was intraperitoneally administered at 48 h, 72 h and 96 h for TCZ, PH1, PH2 and hIL-6 control (IL-6 + vehicle) group. Vehicle group was injected with buffer, instead of antibody and hIL-6 (vehicle + vehicle). The concentration of SAA was measured as a marker of neutralization of hIL-6R²¹. Time profiles of plasma antibody concentration (a) and plasma SAA concentration (b) are shown. Each data point represents the mean \pm s.d. for antibody concentration and mean \pm s.e.m. for SAA concentration ($n = 3$ –4 each). (c,d) *In vivo* study of TCZ and PH2 in normal mice. hsIL-6R, TCZ + hsIL-6R and PH2 + hsIL-6R were intravenously administered at single doses of 50 $\mu\text{g/kg}$ for hsIL-6R and 1 mg/kg for antibody. Time profiles of plasma antibody concentration (c) and plasma hsIL-6R concentration (d) are shown. Each data point represents the mean \pm s.d. ($n = 3$ each).

hsIL-6R in complex with PH2 is not derived from the presence of more free hsIL-6R, but from the pH-dependent dissociation of hsIL-6R in acidic endosomes. In other words, a complex comprising TCZ and sIL-6R would be taken up by endothelial cells in a nonspecific manner, bind to FcRn and be recycled back to plasma as a complex. This would reduce the clearance of sIL-6R and allow a single antigen binding site to bind to only one sIL-6R molecule during its lifetime in plasma (Supplementary Fig. 2c). On the other hand, our study suggests that PH2 dissociates from sIL-6R within the endosome and sIL-6R is then degraded by the lysosome. The antibody, however, is recycled back to plasma by FcRn, and recycled free antibody in plasma can bind to another sIL-6R molecule, thereby enhancing the clearance of sIL-6R and allowing a single antibody to repeatedly neutralize multiple sIL-6R molecules (Supplementary Fig. 2d). These studies on mIL-6R and sIL-6R suggest that engineering a pH-dependent binding of TCZ prolonged the pharmacokinetics and SAA inhibition in hIL-6R transgenic mice by allowing the antibody to dissociate from mIL-6R and sIL-6R in acidic endosomes. In both cases, dissociated free antibody would be recycled back to plasma and reused for binding to another IL-6R, thereby enhancing the turnover of IL-6R binding and prolonging the duration of IL-6R neutralization.

We next compared the *in vivo* effect of engineering a pH-dependent binding with the more common approach of improving antigen binding affinity in a cynomolgus monkey model. We affinity matured TCZ by introducing multiple mutations in the variable region using standard procedure¹⁰. To reduce nonantigen-mediated clearance, TCZ, an affinity-matured variant (AM), and PH2 were further engineered to improve their binding to human FcRn at pH 6.0 (ref. 12). This generated constructs TCZ-FcRn, AM-FcRn and PH2-FcRn, respectively. Owing to the N434A mutation introduced in these Fc-engineered variants, these variants exhibited an approximately three- to fourfold



stronger binding affinity to human FcRn than their counterparts (data not shown). The affinity and pH dependency of the PH2-FcRn to hsIL-6R binding were similar to these features of the parental PH2 molecule. The affinity of the AM-FcRn to hsIL-6R binding was ~100-fold stronger than that of either TCZ or TCZ-FcRn, and AM-FcRn exhibited little pH dependency (Table 1 and Fig. 1d-f). The binding affinities and pH dependencies of binding of these antibodies to cynomolgus monkey sIL-6R was similar to that of hsIL-6R (data not shown). TCZ, TCZ-FcRn, PH2-FcRn and AM-FcRn were intravenously administered to cynomolgus monkeys, each at a dose of 1 mg/kg. Cynomolgus IL-6 was injected, and the abundance of C-reactive protein (CRP) was measured as a pharmacodynamic marker of IL-6R neutralization²³. Time profiles for changes in antibody and CRP plasma concentrations are shown in Figure 3a and Figure 3b, respectively, and the associated pharmacokinetic parameters are shown in Supplementary Table 4. Comparison between TCZ and TCZ-FcRn demonstrated that increasing FcRn binding affinity to reduce non-antigen-mediated clearance improved the pharmacokinetics and duration of IL-6R neutralization by only a small extent. This is most likely because antigen-mediated clearance was the major contribution to the clearance at 1 mg/kg. Notably, compared with TCZ-FcRn, AM-FcRn (which has ~100-fold improved affinity for IL-6R) did not improve the pharmacokinetics, and only a small improvement in the duration of IL-6R neutralization could be achieved. The significant improvement in both pharmacokinetics and duration of IL-6R neutralization could be achieved by the enhanced antigen binding turnover derived from pH-dependent binding in PH2-FcRn. As AM-FcRn has a very high affinity (52 pM) and increasing the binding affinity to FcRn has improved the nonantigen-mediated clearance, the duration of IL-6R neutralization by 1 mg/kg of AM-FcRn (8 d) would be near the maximum for a conventional antibody. Presumably, the amount of IL-6R turned over in 8 d is close to the amount of antigen binding sites of 1 mg/kg of antibody. On the other hand, PH2-FcRn exhibited sustained IL-6R neutralization for 16 d, a level that could not be achieved by a conventional antibody without antigen binding turnover.

IL-6R blockade by TCZ is reported to be efficacious in treating rheumatoid arthritis¹. As subcutaneous injection at a reduced dosing frequency is especially desirable for such chronic diseases, we injected cynomolgus monkeys subcutaneously with TCZ and PH2-FcRn, each at a dose of 2 mg/kg. Time profiles for changes in the antibody and CRP plasma concentrations are shown in Figure 3c and Figure 3d, respectively, and the associated pharmacokinetic parameters are shown in Supplementary Table 5. Whereas 2 mg/kg of TCZ inhibited CRP for ~1 week, 2 mg/kg of PH2-FcRn inhibited CRP for at least 4 weeks. As CRP is an efficient marker for IL-6R blockade by tocilizumab²⁴, this result suggests that PH2-FcRn could provide clinical benefits when administered subcutaneously once a month. Superior pharmacokinetics and improved duration of IL-6R neutralization of PH2-FcRn compared to TCZ were also observed in hIL-6R transgenic mice (Supplementary Fig. 3a-b).

To quantitatively understand the improved pharmacokinetics of PH2-FcRn relative to TCZ in cynomolgus monkey, we fitted the plasma concentration time profiles of TCZ and PH2-FcRn at various i.v. doses to an antigen-mediated clearance and recycling model (Supplementary Fig. 4). The parameters obtained for TCZ and PH2-FcRn are shown in Supplementary Table 6. We determined the relative recycling ratio of the antibody from the endosome to be 0.75 for PH2-FcRn. This indicates that 75% of the internalized antibody is recycled back to plasma as a free antibody compared to TCZ. The nonantigen-mediated elimination rate of antibody was 1.2-fold lower for PH2-FcRn than for TCZ. This difference is largely due to increased binding to FcRn. We propose that a combination of pH-dependent IL-6R binding to reduce the antigen-mediated clearance and increased binding affinity to FcRn to reduce the non-antigen-mediated clearance contributed cooperatively to improve the antibody pharmacokinetics.

Although it remains to be seen how broadly the strategy used in our study can be extended, we believe that engineering pH-dependent antigen binding should improve the duration of antigen neutralization of other therapeutic antibodies. A combination of pH-dependent

binding with affinity maturation and Fc-engineering to increase binding to FcRn should further enhance the clinical potential of many antibody drugs. In this study, pH-dependent binding antibody was generated by histidine scanning. As the dissociation of an antibody from the antigen within the endosome requires destabilization of antibody-antigen interaction upon protonation of the histidine residues in the acidic pH, the feasibility of engineering pH-dependent antigen binding antibody by incorporating histidine residues may depend on electrostatic features of the antigen epitope. Nevertheless, we have successfully applied the same histidine scanning approach to two other antibodies that target different antigens (data not shown). In one of these cases, we achieved pH-dependent binding with picomolar affinity at pH 7.4 (data not shown). These findings suggest the broader potential of the approach to target antigens with high rates of *in vivo* turnover. Until now, these targets have been relatively recalcitrant to efforts based on simply increasing the affinities of therapeutic antibodies for their targets. This approach also seems well-suited to reducing the frequency and/or the size of doses required for currently approved antibody therapeutics.

METHODS

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

Accession codes. GenBank: AB591055-AB591062.

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

ACKNOWLEDGMENTS

We thank T. Kishimoto at the Graduate School of Frontier Biosciences, Osaka University and T. Taga at the Kumamoto University Graduate School of Medical Sciences for kindly providing human IL-6R transgenic mice; colleagues in Chugai Research Institute for Medical Science, Inc., O. Ueda, T. Tachibe, M. Kakefuda and K. Jishage for breeding human IL-6R transgenic mice, T. Matsuura, M. Hiranuma, T. Koike, R. Takemoto, H. Azabu, T. Sakamoto, H. Sano and M. Kawaharada for carrying out *in vivo* experiments, and M. Fujii and A. Maeno for antibody vector construction, expression and purification; and colleagues in Chugai Pharmaceutical Co. Ltd., K. Kasutani, F. Mimoto and K. Esaki for carrying out *in vitro* experiments.

AUTHOR CONTRIBUTIONS

T.I. led the overall pH-dependent binding antibody program, designed experiments, generated tocilizumab variants and wrote the manuscript. S.I. and A.M. generated tocilizumab variants. T.T., R.T., Y.H. and K. Haraya performed *in vivo* studies. S.S. led the anti-IL-6R antibody program. C.M. and A.H. performed affinity analysis of tocilizumab variants. T. Watanabe performed *in vitro* studies of tocilizumab variants. Y.D. and T. Wakabayashi performed purification of tocilizumab variants. S.K. and T.M. provided structural information for designing tocilizumab variants. Y.S., T.K., Y.N., Y.A., Y.K., and K. Hattori provided direction and guidance for the various functional areas.

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/>.

- Mircic, M. & Kavanaugh, A. The clinical efficacy of tocilizumab in rheumatoid arthritis. *Drugs Today (Barc)* **45**, 189–197 (2009).
- Hirota, H., Yoshida, K., Kishimoto, T. & Taga, T. Continuous activation of gp130, a signal-transducing receptor component for interleukin 6-related cytokines, causes myocardial hypertrophy in mice. *Proc. Natl. Acad. Sci. USA* **92**, 4862–4866 (1995).
- Chan, A.C. & Carter, P.J. Therapeutic antibodies for autoimmunity and inflammation. *Nat. Rev. Immunol.* **10**, 301–316 (2010).
- Weiner, L.M., Surana, R. & Wang, S. Monoclonal antibodies: versatile platforms for cancer immunotherapy. *Nat. Rev. Immunol.* **10**, 317–327 (2010).
- Ohsugi, Y. & Kishimoto, T. The recombinant humanized anti-IL-6 receptor antibody tocilizumab, an innovative drug for the treatment of rheumatoid arthritis. *Expert Opin. Biol. Ther.* **8**, 669–681 (2008).
- Tabrizi, M.A., Tseng, C.M. & Roskos, L.K. Elimination mechanisms of therapeutic monoclonal antibodies. *Drug Discov. Today* **11**, 81–88 (2006).
- Tabrizi, M., Bornstein, G.G. & Suria, H. Biodistribution mechanisms of therapeutic monoclonal antibodies in health and disease. *AAPS J.* **12**, 33–43 (2010).
- Ng, C.M., Stefanich, E., Anand, B.S., Fielder, P.J. & Vaickus, L. Pharmacokinetics/pharmacodynamics of nondepleting anti-CD4 monoclonal antibody (TRX1) in healthy human volunteers. *Pharm. Res.* **23**, 95–103 (2006).
- Kelley, S.K. *et al.* Preclinical pharmacokinetics, pharmacodynamics, and activity of a humanized anti-CD40 antibody (SGN-40) in rodents and non-human primates. *Br. J. Pharmacol.* **148**, 1116–1123 (2006).
- Bostrom, J., Lee, C.V., Haber, L. & Fuh, G. Improving antibody binding affinity and specificity for therapeutic development. *Methods Mol. Biol.* **525**, 353–376 (2009).
- Dall'Acqua, W.F., Kiener, P.A. & Wu, H. Properties of human IgG1s engineered for enhanced binding to the neonatal Fc receptor (FcRn). *J. Biol. Chem.* **281**, 23514–23524 (2006).
- Deng, R. *et al.* Pharmacokinetics of humanized monoclonal anti-tumor necrosis factor- α antibody and its neonatal Fc receptor variants in mice and cynomolgus monkeys. *Drug Metab. Dispos.* **38**, 600–605 (2010).
- Zalevsky, J. *et al.* Enhanced antibody half-life improves *in vivo* activity. *Nat. Biotechnol.* **28**, 157–159 (2010).
- Beck, A., Wurch, T., Bailly, C. & Corvaia, N. Strategies and challenges for the next generation of therapeutic antibodies. *Nat. Rev. Immunol.* **10**, 345–352 (2010).
- Rose-John, S., Scheller, J., Elson, G. & Jones, S.A. Interleukin-6 biology is coordinated by membrane-bound and soluble receptors: role in inflammation and cancer. *J. Leukoc. Biol.* **80**, 227–236 (2006).
- Sato, K. *et al.* Reshaping a human antibody to inhibit the interleukin 6-dependent tumor cell growth. *Cancer Res.* **15**, 851–856 (1993).
- Maxfield, F.R. & McGraw, T.E. Endocytic recycling. *Nat. Rev. Mol. Cell Biol.* **5**, 121–132 (2004).
- Sarkar, C.A. *et al.* Rational cytokine design for increased lifetime and enhanced potency using pH-activated “histidine switching”. *Nat. Biotechnol.* **20**, 908–913 (2002).
- Maeda, K., Kato, Y. & Sugiyama, Y. pH-dependent receptor/ligand dissociation as a determining factor for intracellular sorting of ligands for epidermal growth factor receptors in rat hepatocytes. *J. Control. Release* **82**, 71–82 (2002).
- Burmeister, W.P., Huber, A.H. & Bjorkman, P.J. Crystal structure of the complex of rat neonatal Fc receptor with Fc. *Nature* **372**, 379–383 (1994).
- Mihara, M. *et al.* Anti-interleukin 6 receptor antibody inhibits murine AA-amyloidosis. *J. Rheumatol.* **31**, 1132–1138 (2004).
- Finkelman, F.D. *et al.* Anti-cytokine antibodies as carrier proteins. Prolongation of *in vivo* effects of exogenous cytokines by injection of cytokine-anti-cytokine antibody complexes. *J. Immunol.* **151**, 1235–1244 (1993).
- Shinkura, H. *et al.* *In vivo* blocking effects of a humanized antibody to human interleukin-6 receptor on interleukin-6 function in primates. *Anticancer Res.* **18**, 1217–1221 (1998).
- Genovese, M.C. *et al.* Interleukin-6 receptor inhibition with tocilizumab reduces disease activity in rheumatoid arthritis with inadequate response to disease-modifying antirheumatic drugs: the tocilizumab in combination with traditional disease-modifying antirheumatic drug therapy study. *Arthritis Rheum.* **58**, 2968–2980 (2008).

ONLINE METHODS

Generation of TCZ variants with pH-dependent binding. TCZ variants with histidine mutation in CDR or framework residues were generated by site-directed mutagenesis. SPR measurements of histidine-mutated variants were performed to evaluate their ability to bind to hIL-6R pH-dependently. Various mutations were combined to generate PH1 and PH2 with the desired pH-dependent binding. An affinity-improved variant of tocilizumab (AM) was generated using the standard affinity maturation procedure¹⁰. TCZ, PH2 and AM were Fc-engineered for improved binding to FcRn¹² to generate TCZ-FcRn, PH2-FcRn and AM-FcRn, respectively. The antibodies were expressed either stably or transiently in Chinese hamster ovary (CHO) or HEK293 cells and purified.

TCZ, PH1, PH2, TCZ-FcRn, PH2-FcRn and AM-FcRn were assessed for their binding affinity to recombinant hIL-6R²⁵ at pH 7.4 and pH 6.0, at 37 °C using Biacore T100 (GE Healthcare). Each antibody was captured onto the anti-human IgG (γ-chain specific) F(ab')₂ fragment (Sigma-Aldrich) immobilized CM5 sensor chip, then hIL-6R was injected over the flow cell. Kinetic binding constants were determined through global fitting using Biacore T100 Evaluation Software, version 2.0.2 (GE Healthcare).

In vivo study of TCZ, PH1 and PH2 in hIL-6R transgenic mice. 25 mg/kg doses of TCZ, PH1 and PH2 were administered to hIL-6R transgenic mice by single i.v. injection. Extent of hIL-6R neutralization was assessed by i.p. injection of 4 μg/kg hIL-6 (TORAY) at various time points. 20 mg/kg of rat anti-mouse IL-6R antibody, MR16-1 (ref. 26), was injected before hIL-6 injection to inhibit the effect of hIL-6 binding to mouse IL-6R. Blood samples were collected 6 h after the injection of hIL-6. All animal experiments in this study were performed in accordance with the Guidelines for the Care and Use of Laboratory Animals at Chugai Pharmaceutical Co., Ltd.

To determine the plasma antibody concentrations, enzyme-linked immunosorbent assay (ELISA) 96-well plates were pre-coated with anti-human IgG (γ-chain specific) F(ab')₂ fragment of antibody (Sigma) and appropriately blocked. Antibody was detected by biotinylated goat anti-human IgG antibody (Southern Biotechnology Associates) and streptavidin-alkaline phosphatase conjugate (Roche Diagnostics) using alkaline phosphatase substrate. Pharmacokinetic parameters were calculated using noncompartmental analysis of WinNonlin Professional (version 4.0.1) software (Pharsight). Time-points for T_{1/2} calculation were automatically set by WinNonlin on the basis of the result of the time profile in plasma concentrations. Plasma SAA concentrations were determined 6 h after administration of hIL-6 using Mouse SAA ELISA kit (Invitrogen).

In vivo study of TCZ and PH2 in normal mice. To evaluate the effect of pH-dependent binding on mIL-6R and sIL-6R, hIL-6R alone or mixed with excess amount of antibody was administered to C57BL/6J normal mice (Charles River) by single i.v. injection. Each group received 50 μg/kg hIL-6R, but the second and third groups additionally received 1 mg/kg TCZ or 1 mg/kg PH2, respectively. hIL-6R and antibody were mixed before administration.

Plasma antibody concentrations were determined using anti-human IgG (γ-chain specific) F(ab')₂ fragment of antibody coated on ELISA 96-well plates, and detected by hIL-6R, biotinylated anti-hIL-6R antibody and

Streptavidin-PolyHRP80 (Stereospecific Detection Technologies) using peroxidase substrate. Pharmacokinetic parameters were calculated as described above.

To determine plasma hIL-6R concentration, plasma samples were mixed with SULFO-TAG (Meso Scale Discovery)-labeled anti-hIL-6R antibody (R&D Systems), biotinylated anti-hIL-6R antibody (R&D Systems), and excess amount of tocilizumab (TCZ) and dispensed into the streptavidin-coated standard 96-well plate (Meso Scale Discovery). After incubation and washing, read Buffer T (× 4) (Meso Scale Discovery) was added and measured by SECTOR PR 400 reader (Meso Scale Discovery). The plasma hIL-6R concentrations were determined from the standard curve and pharmacokinetic parameters were calculated as described above.

In vivo study of TCZ, TCZ-FcRn, PH2-FcRn and AM-FcRn in cynomolgus monkeys. In the i.v. study, a single dose of 1 mg/kg of TCZ or TCZ-FcRn or PH2-FcRn or AM-FcRn was intravenously administered, and in the subcutaneous (s.c.) study, a single dose of 2 mg/kg of TCZ or PH2-FcRn was subcutaneously administered to cynomolgus monkeys. To assess the extent of IL-6R neutralization, 5 μg/kg of recombinant cynomolgus IL-6 (Chugai Pharmaceutical) was administered subcutaneously in the lower back at various time points. Blood samples were collected at an appropriate time after each injection.

Plasma antibody concentrations of the i.v. study were determined as described above for normal mice using recombinant cynomolgus sIL-6R (Chugai Pharmaceutical) instead of hIL-6R. Plasma TCZ or PH2-FcRn concentrations of the s.c. study were determined using hIL-6R-coated ELISA 96-well plates. TCZ was detected by biotinylated goat anti-human IgG antibody and streptavidin-alkaline phosphatase conjugate using alkaline phosphatase substrate. PH2-FcRn was detected by biotinylated anti-PH2-FcRn antibody and Streptavidin-PolyHRP80 using peroxidase substrate. Pharmacokinetic parameters were calculated as described above.

Plasma CRP concentrations were determined 24 h after administration of cynomolgus IL-6 using Cias R CRP (Kanto Chemical).

Pharmacokinetic analysis using antigen-mediated clearance and recycling model. The plasma concentration time profiles of TCZ at doses of 0.5, 1.0, 2.0, 5.0 and 50 mg/kg (i.v.) in cynomolgus monkey were fitted to an antigen-mediated clearance and recycling model (Supplementary Fig. 4) and V_d (volume of distribution), k_{el} (elimination rate constant), k_{12} (transfer rate constant from central to peripheral compartment), k_{21} (transfer rate constant from peripheral to central compartment) and R_{total} (total amount of mIL-6R) were optimized assuming F_{rec} (fraction recycled from endosome) of TCZ to be 0. Then, the plasma concentration time profiles of PH2-FcRn at doses of 0.5, 1.0, 2.0 and 5.0 mg/kg (i.v.) were fitted to the same model and V_d , k_{el} , k_{12} , k_{21} and F_{rec} were optimized using fixed R_{total} value derived from TCZ data. The k_a and k_d values were from SPR data (Table 1). The k_{int} (internalization rate constant) value was from relevant literature²⁷.

25. Yasukawa, K. *et al.* Purification and characterization of soluble human IL-6 receptor expressed in CHO cells. *J. Biochem.* **108**, 673–676 (1990).

26. Okazaki, M., Yamada, Y., Nishimoto, N., Yoshizaki, K. & Mihara, M. Characterization of anti-mouse interleukin-6 receptor antibody. *Immunol. Lett.* **84**, 231–240 (2002).

27. Gerhartz, C. *et al.* Biosynthesis and half-life of the interleukin-6 receptor and its signal transducer gp130. *Eur. J. Biochem.* **223**, 265–274 (1994).

Programmable *in situ* amplification for multiplexed imaging of mRNA expression

Harry M T Choi¹, Joann Y Chang¹, Le A Trinh², Jennifer E Padilla¹, Scott E Fraser^{1,2} & Niles A Pierce^{1,3}

***In situ* hybridization methods enable the mapping of mRNA expression within intact biological samples^{1,2}. With current approaches, it is challenging to simultaneously map multiple target mRNAs within whole-mount vertebrate embryos^{3–6}, representing a significant limitation in attempting to study interacting regulatory elements in systems most relevant to human development and disease. Here, we report a multiplexed fluorescent *in situ* hybridization method based on orthogonal amplification with hybridization chain reactions (HCR)⁷. With this approach, RNA probes complementary to mRNA targets trigger chain reactions in which fluorophore-labeled RNA hairpins self-assemble into tethered fluorescent amplification polymers. The programmability and sequence specificity of these amplification cascades enable multiple HCR amplifiers to operate orthogonally at the same time in the same sample. Robust performance is achieved when imaging five target mRNAs simultaneously in fixed whole-mount and sectioned zebrafish embryos. HCR amplifiers exhibit deep sample penetration, high signal-to-background ratios and sharp signal localization.**

Each cell in a multicellular organism contains the same genome, yet the regulatory circuits encoded within this genome implement a developmental program yielding significant spatial heterogeneity and complexity. *In situ* hybridization methods are an essential tool for elucidating developmental and pathological processes, enabling imaging of mRNA expression in a morphological context from subcellular to organismal-length scales^{1,2,8–21}.

Due to variability between biological specimens, the accurate mapping of spatial relationships between regulatory loci of different genes requires multiplexed experiments in which multiple mRNAs are imaged in a single biological sample. Within intact vertebrate embryos, enzymatic *in situ* amplification methods based on catalytic deposition of reporter molecules are currently the method of choice to achieve high signal-to-background ratios^{4,5,22,23}. The key difficulty is the lack of orthogonal deposition chemistries, necessitating serial multiplexing approaches in which two^{3,5} or three^{4,6} target mRNAs are detected in succession using cumbersome procedures that progressively degrade the sample as the number of target mRNAs increases. Here, we overcome this difficulty by programming orthogonal HCR amplifiers⁷ that function as independent molecular instruments,

simultaneously reading out the expression patterns of five target mRNAs from within a single intact biological sample.

An HCR amplifier consists of two nucleic acid hairpin species (H1 and H2 in Fig. 1a) that are designed to co-exist metastably in the absence of a nucleic acid initiator (I)⁷. Each HCR hairpin consists of an input domain with an exposed single-stranded toehold and an output domain with a single-stranded toehold sequestered in the hairpin loop. Hybridization of the initiator to the input domain of H1 (Fig. 1a, '1-2') opens the hairpin to expose its output domain (Fig. 1a, '3*-2*'). Hybridization of this output domain to the input domain of H2 (Fig. 1a, '2-3') opens the hairpin to expose an output domain (Fig. 1a, '2*-1*') identical in sequence to the initiator. Regeneration of the initiator sequence provides the basis for a chain reaction of alternating H1 and H2 polymerization steps leading to formation of a nicked double-stranded 'polymer'. If the initiator is absent, the hairpins are metastable (that is, kinetically impeded from polymerizing) due to the sequestration of the output toeholds in the hairpin loops.

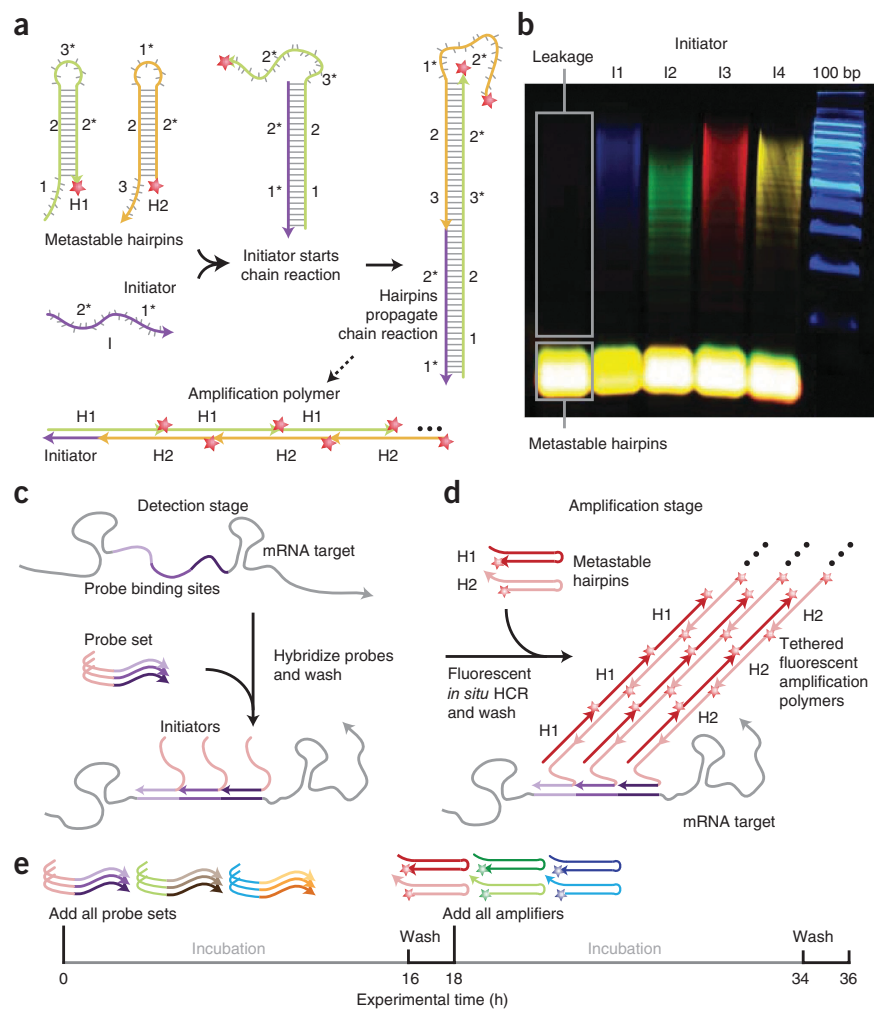
This mechanism has two properties that are important when attempting to achieve simultaneous multiplexed *in situ* amplification in vertebrate embryos. First, the programmable chemistry of nucleic acid base pairing suggests the feasibility of engineering orthogonal HCR amplifiers that operate independently in the same sample at the same time. Second, in contrast to molecular self-assembly by means of traditional annealing protocols in which components interact as soon as they are mixed together²⁴, HCR is an isothermal triggered self-assembly process. Hence, hairpins should penetrate the sample before undergoing triggered self-assembly *in situ*, suggesting the potential for deep sample penetration and high signal-to-background ratios.

Despite previous successes in implementing HCR in a test tube^{7,25}, it proved challenging to engineer HCR hairpins for *in situ* hybridization due to the stringent hybridization conditions that are required to destabilize nonspecific binding (40% hybridization buffer; **Supplementary Notes**). The free energy of each HCR polymerization step arises from the enthalpic benefit of forming additional stacked base pairs between the toehold in the output domain at the growing end of the polymer and the toehold in the input domain of a newly recruited hairpin, as well as from the entropic benefit of opening the hairpin loop of the recruited hairpin. The original HCR system used DNA hairpins with 6-nt toeholds and loops and 18-bp stems⁷ (resulting in six stacked base pairs plus the opening of a 6-nt hairpin loop per polymerization step). Preliminary test tube and *in situ* hybridization studies revealed

¹Department of Bioengineering, California Institute of Technology, Pasadena, California, USA. ²Department of Biology, California Institute of Technology, Pasadena, California, USA. ³Department of Applied & Computational Mathematics, California Institute of Technology, Pasadena, California, USA. Correspondence should be addressed to N.A.P. (niles@caltech.edu).

Received 28 June; accepted 24 September; published online 31 October 2010; doi:10.1038/nbt.1692

Figure 1 Multiplexed *in situ* hybridization using fluorescent HCR *in situ* amplification. (a) HCR mechanism. Metastable fluorescent RNA hairpins self-assemble into fluorescent amplification polymers upon detection of a specific RNA initiator. Initiator I nucleates with hairpin H1 via base pairing to single-stranded toehold '1', mediating a branch migration³⁰ that opens the hairpin to form complex I-H1 containing single-stranded segment '3*-2*'. This complex nucleates with hairpin H2 by means of base pairing to toehold '3', mediating a branch migration that opens the hairpin to form complex I-H1-H2 containing single-stranded segment '2*-1*'. Thus, the initiator sequence is regenerated, providing the basis for a chain reaction of alternating H1 and H2 polymerization steps. Red stars denote fluorophores. (b) Validation in a test tube. Agarose gel demonstrating orthogonal amplification in a reaction volume containing four HCR amplifiers and zebrafish total RNA. Minimal leakage from metastable states is observed in the absence of initiators. (c) Detection stage. Probe sets are hybridized to mRNA targets and then unused probes are washed from the sample. (d) Amplification stage. Initiators trigger self-assembly of tethered fluorescent amplification polymers and then unused hairpins are washed from the sample. (e) Experimental timeline. The same two-stage protocol is used independent of the number of target mRNAs. For multiplexed experiments (three-color example depicted), probe sets for different target mRNAs carry orthogonal initiators that trigger orthogonal HCR amplification cascades labeled by spectrally distinct fluorophores.



that this small-loop DNA-HCR system did not polymerize under stringent hybridization conditions due to insufficient free energy per polymerization step²⁶.

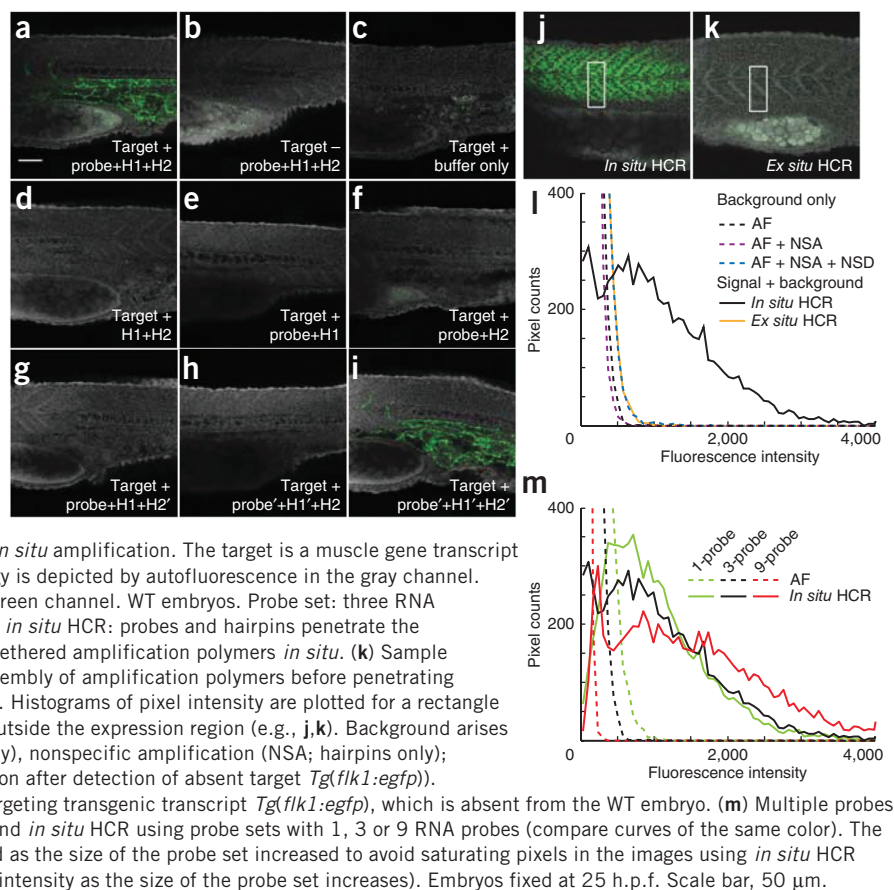
Thus, we confronted the challenge of engineering new HCR hairpins that retain two key properties under these conditions: (i) hairpin metastability in the absence of the initiator, (ii) hairpin polymerization in the presence of the initiator. Previous experience suggested that these two objectives are at odds. Hairpin metastability is promoted by reducing toehold and loop size; hairpin polymerization is promoted by increasing toehold and loop size.

Secondary structure free energy parameters have not been measured for stringent hybridization conditions, so we could not resize components based on computational simulation. Instead, we used test tube and *in situ* control experiments to measure the minimum hairpin toehold and loop length necessary for stable hybridization. Imposing this design constraint to promote hairpin polymerization did not prevent us from retaining hairpin metastability under the same stringent hybridization conditions. To partially counteract the necessary increase in hairpin size, we switched from DNA to RNA hairpins to exploit the enhanced stability of stacked RNA base pairs relative to DNA base pairs. The resulting big-loop RNA-HCR system has 10-nt toeholds and loops and 16-bp stems. The results of the test tube study presented in **Figure 1b** illustrate four HCR amplifiers operating simultaneously and orthogonally in a background of zebrafish total RNA under stringent hybridization conditions. The hairpins exhibit metastability in the absence of initiators; the introduction of a single initiator species selectively triggers the cognate polymerization reaction.

We perform *in situ* hybridization in two stages independent of the number of target mRNAs (**Fig. 1c–e**). In the detection stage, all target mRNAs are detected simultaneously via *in situ* hybridization of complementary RNA probes; unused probes are washed from the sample. Each target mRNA is addressed by a probe set comprising one or more RNA probe species carrying identical initiators; different targets are addressed by probe sets carrying orthogonal initiators. In the amplification stage, optical readouts are generated for all target mRNAs simultaneously using fluorescent *in situ* HCR. Orthogonal initiators trigger orthogonal hybridization chain reactions in which metastable RNA hairpins self-assemble into tethered amplification polymers labeled with spectrally distinct fluorophores; unused hairpins are washed from the sample before imaging.

To validate HCR *in situ* amplification in fixed whole-mount zebrafish embryos, we first targeted a transgenic mRNA, observing bright staining with the expected expression pattern (**Fig. 2a**). Wild-type embryos (lacking the target) show minimal staining (**Fig. 2b**), comparable to the autofluorescence observed in the absence of probes and hairpins (**Fig. 2c**). As expected, amplification is not observed if the probe or either of the two hairpin species is omitted (**Fig. 2d–f**). To verify that the staining in **Figure 2a** results from the intended polymerization mechanism rather than from aggregation of closed hairpins, alteration of one or both hairpin stem sequences yields the expected loss (**Fig. 2g,h**) and recovery (**Fig. 2i**) of signal.

Figure 2 Validation of fluorescent HCR *in situ* amplification in fixed whole-mount zebrafish embryos. (a–i) The target is the transgenic transcript *Tg(flk1:egfp)*, expressed below the notochord and between the somites (see the expression atlas of Fig. 3a). Embryo morphology is depicted by autofluorescence in the gray channel. Probe set: 1 RNA probe. Fluorescent staining (green channel) using *in situ* HCR in Target⁺ (a) and Target⁻ (b) embryos compared to (green channel) autofluorescence in the absence of probes and hairpins (c). No amplification in the absence of probes (d) or of one hairpin species (e,f). Modification of hairpin stem sequences (H1', H2') disrupts (g,h) and restores (i) toehold-mediated branch migration, confirming that staining arises from triggered polymerization rather than from random aggregation of hairpins. Typical for zebrafish, the yolk sack (bottom left of each panel) often exhibits autofluorescence. (j–m) Characterizing the signal-to-background ratio for fluorescent HCR *in situ* amplification. The target is a muscle gene transcript (*desm*) expressed in the somites. Embryo morphology is depicted by autofluorescence in the gray channel. Pixel intensity histograms are calculated using the green channel. WT embryos. Probe set: three RNA probes, except panel m. (j) Sample penetration with *in situ* HCR: probes and hairpins penetrate the sample before executing triggered self-assembly of tethered amplification polymers *in situ*. (k) Sample penetration with *ex situ* HCR: probes trigger self-assembly of amplification polymers before penetrating the sample. (l) Background and signal contributions. Histograms of pixel intensity are plotted for a rectangle partially within the expression region and partially outside the expression region (e.g., j,k). Background arises from three sources: autofluorescence (AF; buffer only), nonspecific amplification (NSA; hairpins only); nonspecific detection (NSD; *in situ* HCR amplification after detection of absent target *Tg(flk1:egfp)*). NSD studies use a probe set of three RNA probes targeting transgenic transcript *Tg(flk1:egfp)*, which is absent from the WT embryo. (m) Multiple probes per mRNA target. Comparison of autofluorescence and *in situ* HCR using probe sets with 1, 3 or 9 RNA probes (compare curves of the same color). The microscope photomultiplier tube gain was decreased as the size of the probe set increased to avoid saturating pixels in the images using *in situ* HCR amplification (this accounts for the reduction in AF intensity as the size of the probe set increases). Embryos fixed at 25 h.p.f. Scale bar, 50 μ m.



Detection and amplification components must successfully penetrate an embryo to generate signal at the site of an mRNA target. HCR is a triggered self-assembly mechanism, offering the conceptual benefit that small RNA probes and hairpins penetrate the embryo before generating larger, less-mobile amplification polymers at the site of mRNA targets. To assess the practical significance of these properties, we imaged an endogenous mRNA with a superficial expression pattern, comparing *in situ* HCR to the *ex situ* HCR alternative in which amplification polymers are pre-assembled before penetrating the sample. The images of Figure 2j,k and the pixel-intensity histograms of Figure 2l demonstrate dramatic signal loss using *ex situ* HCR. This result is consistent with the general experience that large, multilabeled probes suffer from reduced sample penetration and confirms that it is desirable to penetrate the sample with small components that self-assemble in a triggered fashion at the site of mRNA targets.

In situ amplification is intended to generate a high signal-to-background ratio to enable accurate mapping of mRNA expression patterns. With our approach, signal is produced when specifically hybridized probes initiate specific HCR amplification to yield fluorescent polymers tethered to cognate mRNA targets. Background can arise from three sources: nonspecific detection (probes that bind nonspecifically and are subsequently amplified), nonspecific amplification (hairpins and polymers that are not hybridized to cognate initiators) and autofluorescence (inherent fluorescence of the fixed embryo). To characterize the relative magnitudes of these effects, we imaged an mRNA target with a sharply defined region of expression and plotted histograms of pixel intensity within a rectangle that crosses the boundary of this expression region. The pixel intensity histograms of Figure 2l reveal that autofluorescence is the primary source of background, that nonspecific

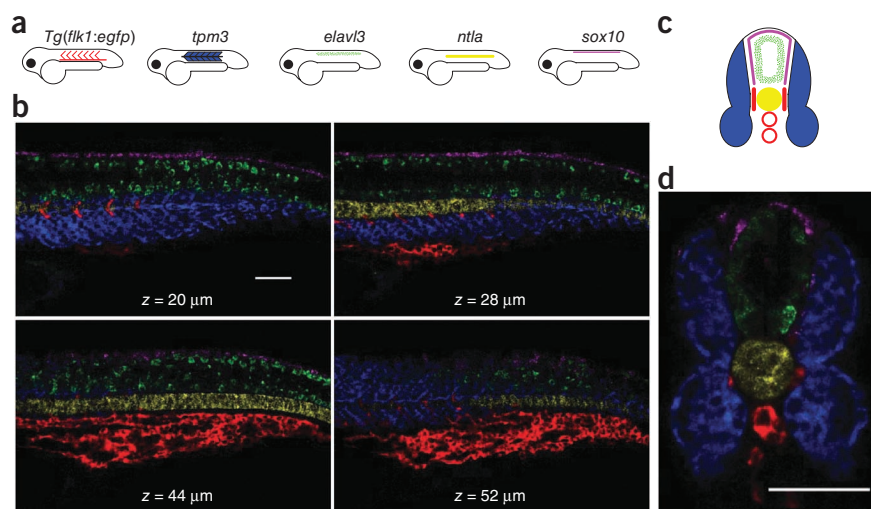
detection contributes a small amount of additional background and that nonspecific amplification contributes negligibly to background. By comparison, the signal generated using *in situ* HCR amplification yields pixel intensities that are significantly higher than background.

The observation that autofluorescence is the dominant source of background suggests that addressing each target mRNA with a probe set comprising multiple probes^{13,19} would further increase the signal-to-background ratio. Subsequent HCR *in situ* amplification would then decorate each target with an array of amplification polymers. Figure 2m demonstrates that the ratio of signal to autofluorescence increases with the number of probes per target. Notably, using *in situ* HCR, the pixel intensity distribution is bimodal using either three or nine probes per target, with a peak at low intensity corresponding to background (from the portion of the rectangle outside the expression region) and a broad distribution at higher intensities corresponding to signal (from the portion of the rectangle within the expression region). High signal-to-background is demonstrated for a target mRNA with a lower level of expression in Supplementary Notes.

The fundamental benefit of using orthogonal HCR amplifiers is the ability to perform simultaneous *in situ* amplification for multiple target mRNAs, enabling straightforward multiplexing. Figure 3 demonstrates simultaneous imaging of five target mRNAs in fixed whole-mount and cross-sectioned zebrafish embryos. Targets were detected using five probe sets carrying five orthogonal initiators and amplification was performed using five orthogonal HCR amplifiers carrying five spectrally distinct fluorophores. The expression patterns in the cross-sectioned embryo confirm that HCR signal survives vibratome sectioning.

Using HCR *in situ* amplification, each amplification polymer is expected to remain tethered to its initiating probe, suggesting the

Figure 3 Multiplexed imaging in fixed whole-mount and cross-sectioned zebrafish embryos. (a) Expression atlas for five target mRNAs (lateral view: *Tg(flk1:egfp)*, *tpm3*, *elavl3*, *ntla*, *sox10*). (b) mRNA expression imaged using confocal microscopy at four planes within an embryo. This multiplexed experiment is performed using the same two-stage protocol that is used for single-color experiments (summarized in Fig. 1c–e). Detection is performed using five probe sets carrying orthogonal initiators. The probe sets have different numbers of RNA probes (10, 7, 18, 30, 20) based on the strength of expression of each mRNA target and the strength of the autofluorescence in each channel. Amplification is performed using five orthogonal HCR amplifiers carrying spectrally distinct fluorophores. (c) Expression atlas for five target mRNAs (anterior view). (d) mRNA expression imaged within a 200- μm zebrafish section using confocal microscopy. Vibratome sectioning was performed after HCR *in situ* amplification and post-fixation. See also the image stacks of **Supplementary Movies 1** and **2**. Embryos fixed at 27 h.p.f. Scale bars, 50 μm .



potential for accurate signal localization and co-localization. Here, we test signal localization and co-localization using a four-color, two-target experiment in which one target mRNA is expressed predominantly in the somites and the other is expressed predominantly in the interstices of somites. The two target mRNAs are each detected using two independent probe sets and each of the four probe sets is amplified using a spectrally distinct HCR amplifier. Double detection of a single target mRNA provides a rigorous test of signal co-localization independent of the expression pattern of the target. **Figure 4a,b** reveals sharp co-localization of two signals for each of the two target mRNAs.

Simultaneous mapping of two targets expressed in contiguous cells provides a further test of signal localization. **Figure 4c** demonstrates interleaving of two sharp expression patterns, revealing that the interstitial expression pattern between somites is only the width of a single stretched cell. This study suggests that HCR polymers remain tethered to their initiating probes and demonstrates sharp signal localization and co-localization at the level of single cells within whole-mount zebrafish embryos.

The sequencing of numerous genomes has launched a new era in biology, enabling powerful comparative approaches and revealing the nucleotide sequences that contribute to the differences between species, between individuals of the same species and between cells within an individual. However, knowledge of these sequences is not sufficient to reveal the architecture and function of the biological circuits that account for these differences. Much work remains to

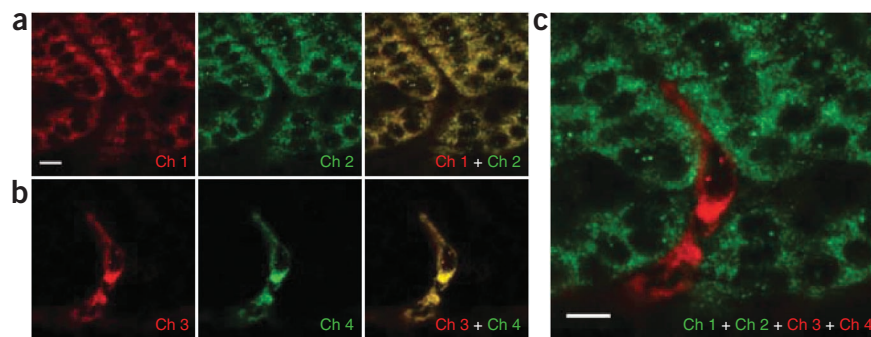
elucidate both the details and the principles of the molecular circuits that regulate development, maintenance, repair and disease within living organisms.

Over four decades⁸, *in situ* hybridization methods have become an indispensable tool for the study of genetic regulation in a morphological context. Current methods of choice for performing enzymatic *in situ* amplification in vertebrate embryos require serial amplification for multiplexed studies^{3–6,22,23}. This shortcoming is a major impediment to the study of interacting regulatory elements *in situ*. For example, simultaneous mapping of three target mRNAs in whole-mount chick embryos requires 5 d using serial *in situ* amplification approaches^{4,6}.

In recent years, researchers in the field of nucleic acid nanotechnology have made much progress in designing nucleic acid molecules that interact and change conformation to execute diverse dynamic functions^{27–29}. Here, we exploit design principles drawn from this experience to engineer small conditional RNAs that interact and change conformation to amplify the expression patterns of multiple target mRNAs in parallel within intact vertebrate embryos. The resulting programmable molecular technology addresses a longstanding need in the biological sciences.

HCR *in situ* amplification enables simultaneous mapping of five target mRNAs in fixed whole-mount and sectioned zebrafish embryos. The programmability and sequence specificity of the HCR mechanism enable all five amplifiers to operate orthogonally in the same sample at the same time. Hence, the time required to map five targets is the

Figure 4 Sharp signal localization and co-localization in fixed whole-mount zebrafish embryos. Redundant two-color mapping of one target mRNA expressed predominantly in the somites (*desm*; two probe sets, two HCR amplifiers, channels 1 and 2) simultaneous with redundant two-color mapping of a second target mRNA expressed predominantly in the interstices of somites (*Tg(flk1:egfp)*; two probe sets, two HCR amplifiers, channels 3 and 4). (a) Sharp co-localization of *desm* signal (Pearson correlation coefficient, $r = 0.93$). (b) Sharp co-localization of *Tg(flk1:egfp)* signal (Pearson correlation coefficient, $r = 0.97$). (c) Sharp signal localization within the two interleaved expression regions. The interstice between somites is only the width of a single stretched cell. Embryos fixed at 27 h.p.f. Scale bars, 10 μm .



same as that required to map one target and the sample degradation that accompanies sequential detection of multiple mRNAs is avoided. We observe that autofluorescence, rather than nonspecific detection or nonspecific amplification, is the dominant source of background in zebrafish. Consequently, the signal-to-background ratio is enhanced by using probe sets with multiple probes, each carrying an HCR initiator. Small fluorophore-labeled amplification components penetrate the sample before undergoing triggered self-assembly to form fluorescent amplification polymers that remain tethered to their initiating probes. The triggered self-assembly property leads to a high signal-to-background ratio and deep sample penetration. The tethering property leads to sharp signal localization and co-localization at the level of single cells within whole-mount zebrafish embryos.

Our approach is potentially suited for use in a variety of biological contexts including fixed cells, embryos, tissue sections and microbial populations. By coupling HCR initiators to aptamer or antibody probes, HCR amplification is also potentially suitable for extension to multiplexed imaging of small molecules and proteins. Further work is required to explore these possibilities.

The HCR amplifiers presented here are suitable for use with diverse mRNA targets because the initiator sequences (and consequently the HCR hairpins) are independent of the mRNA target sequences. Imaging a new target mRNA requires only a new probe set with each probe carrying an HCR initiator.

METHODS

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

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

ACKNOWLEDGMENTS

We thank V.A. Beck, J.S. Bois, S. Venkataraman, J.R. Vieregg and P. Yin for discussions. We thank C. Johnson and A.J. Ewald for performing preliminary studies. We thank J.N. Zadeh for the use of unpublished software. We thank the Caltech Biological Imaging Center and A. Collazo of the House Ear Institute for the use of multispectral confocal microscopes. This work was funded by the US National Institutes of Health (R01 EB006192 and P50 HG004071), the National Science Foundation (CCF-0448835 and CCF-0832824) and the Beckman Institute at Caltech.

AUTHOR CONTRIBUTIONS

S.E.F. and N.A.P. conceived the application of HCR to multiplexed bioimaging; H.M.T.C., J.Y.C., J.E.P. and N.A.P. engineered HCR hairpins for use in stringent hybridization buffers; H.M.T.C. and N.A.P. designed the experiments; H.M.T.C. performed the experiments; L.A.T. selected targets, provided technical guidance and performed the control experiments using traditional *in situ* hybridization; H.M.T.C., L.A.T., S.E.F. and N.A.P. analyzed the data; H.M.T.C. and N.A.P. wrote the manuscript; and all authors edited 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. Qian, X., Jin, L. & Lloyd, R.V. *In situ* hybridization: basic approaches and recent development. *J. Histochem. Technol.* **27**, 53–67 (2004).

2. Silverman, A. & Kool, E. Oligonucleotide probes for RNA-targeted fluorescence *in situ* hybridization. *Adv. Clin. Chem.* **43**, 79–115 (2007).
3. Thisse, B. *et al.* Spatial and temporal expression of the zebrafish genome by large-scale *in situ* hybridization screening. in *Zebrafish: Genetics, Genomics and Informatics*, vol. 77, edn. 2 (eds. Detrich, H.W., Zon, L.I. & Westerfield, M.) 505–519 (Elsevier, 2004).
4. Denkers, N., Garcia-Villalba, P., Rodesch, C.K., Nielson, K.R. & Mauch, T.J. FISHing for chick genes: Triple-label whole-mount fluorescence *in situ* hybridization detects simultaneous and overlapping gene expression in avian embryos. *Dev. Dyn.* **229**, 651–657 (2004).
5. Barroso-Chinea, P. *et al.* Detection of two different mRNAs in a single section by dual *in situ* hybridization: A comparison between colorimetric and fluorescent detection. *J. Neurosci. Methods* **162**, 119–128 (2007).
6. Aclouque, H., Wilkinson, D.G. & Nieto, M.A. *In situ* hybridization analysis of chick embryos in whole-mount and tissue sections. in *Avian Embryology, 2nd Edition* vol. 87 (ed. Bronner-Fraser, M.) 169–185 (Elsevier, 2008).
7. Dirks, R.M. & Pierce, N.A. Triggered amplification by hybridization chain reaction. *Proc. Natl. Acad. Sci. USA* **101**, 15275–15278 (2004).
8. Gall, J.G. & Pardue, M.L. Formation and detection of RNA-DNA hybrid molecules in cytological preparations. *Proc. Natl. Acad. Sci. USA* **63**, 378–383 (1969).
9. Lawrence, J.B., Singer, R.H. & Marselle, L.M. Highly localized tracks of specific transcripts within interphase nuclei visualized by *in situ* hybridization. *Cell* **57**, 493–502 (1989).
10. Tautz, D. & Pfeifle, C. A non-radioactive *in situ* hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene hunchback. *Chromosoma* **98**, 81–85 (1989).
11. Kislaukis, E.H., Li, Z., Singer, R.H. & Taneja, K.L. Isoform-specific 3'-untranslated sequences sort α -cardiac and β -cytoplasmic actin messenger RNAs to different cytoplasmic compartments. *J. Cell. Biol.* **123**, 165–172 (1993).
12. O'Neill, J.W. & Bier, E. Double-label *in situ* hybridization using biotin digoxigenin-tagged RNA probes. *Biotechniques* **17**, 870–875 (1994).
13. Femino, A.M., Fay, F.S., Fogarty, K. & Singer, R.H. Visualization of single RNA transcripts *in situ*. *Science* **280**, 585–590 (1998).
14. Zaidi, A.U., Enomoto, H., Milbrandt, J. & Roth, K.A. Dual fluorescent *in situ* hybridization and immunohistochemical detection with tyramide signal amplification. *J. Histochem. Cytochem.* **48**, 1369–1375 (2000).
15. Player, A.N., Shen, L.-P., Kenny, D., Antao, V.P. & Kolberg, J.A. Single-copy gene detection using branched DNA (bDNA) *in situ* hybridization. *J. Histochem. Cytochem.* **49**, 603–612 (2001).
16. Levsky, J.M., Shenoy, S.M., Pezo, R.C. & Singer, R.H. Single-cell gene expression profiling. *Science* **297**, 836–840 (2002).
17. Kosman, D. *et al.* Multiplex detection of RNA expression in *Drosophila* embryos. *Science* **305**, 846 (2004).
18. Lambros, M.B.K., Natrajan, R. & Reis-Filho, J.S. Chromogenic and fluorescent *in situ* hybridization in breast cancer. *Hum. Pathol.* **38**, 1105–1122 (2007).
19. Raj, A., van den Boggaard, P., Rifkin, S.A., van Oudenaarden, A. & Tyagi, S. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods* **5**, 877–879 (2008).
20. Amann, R. & Fuchs, B.M. Single-cell identification in microbial communities by improved fluorescence *in situ* hybridization techniques. *Nat. Rev. Microbiol.* **6**, 339–348 (2008).
21. Larsson, C., Grundberg, I., Soderberg, O. & Nilsson, M. *In situ* detection and genotyping of individual mRNA molecules. *Nat. Methods* **7**, 395–397 (2010).
22. Harland, R.M. *In situ* hybridization: an improved whole-mount method for *Xenopus* embryos. *Methods Cell Biol.* **36**, 685–695 (1991).
23. Speel, E.J.M., Hopman, A.H.N. & Komminoth, P. Amplification methods to increase the sensitivity of *in situ* hybridization: Play CARD(S). *J. Histochem. Cytochem.* **47**, 281–288 (1999).
24. Feldkamp, U. & Niemeyer, C.M. Rational design of DNA nanoarchitectures. *Angew. Chem. Int. Ed.* **45**, 1856–1876 (2006).
25. Venkataraman, S., Dirks, R.M., Rothemund, P.W.K., Winfree, E. & Pierce, N.A. An autonomous polymerization motor powered by DNA hybridization. *Nat. Nanotechnol.* **2**, 490–494 (2007).
26. Choi, H.M.T. *Programmable In Situ Amplification for Multiplexed Bioimaging*. PhD thesis, California Institute of Technology (2009).
27. Simmel, F.C. & Dittmer, W.U. DNA nanodevices. *Small* **1**, 284–299 (2005).
28. Bath, J. & Turberfield, A.J. DNA nanomachines. *Nat. Nanotechnol.* **2**, 275–284 (2007).
29. Feldkamp, U. & Niemeyer, C.M. Rational engineering of dynamic DNA systems. *Angew. Chem. Int. Ed.* **47**, 3871–3873 (2008).
30. Yurke, B., Turberfield, A.J., Mills, J.A.P., Simmel, F.C. & Neumann, J.L. A DNA-fuelled molecular machine made of DNA. *Nature* **406**, 605–608 (2000).

ONLINE METHODS

Probe synthesis. RNA probes are 81-nt long (26-nt initiator, 5-nt spacer, 50-nt mRNA recognition sequence). mRNAs are addressed by probe sets containing one or more probes that hybridize adjacently at 50-nt binding sites. Probe sequences are displayed in **Supplementary Notes**. RNA probes were synthesized by *in vitro* transcription. The coding strand for each probe contained three random nucleotides and a 19-nt SP6 promoter sequence upstream of the 81-nt initiator-linker-probe sequence. Complementary DNA coding and template strands were ordered (unpurified) from Integrated DNA Technologies (IDT). Strands were resuspended in ultrapure water (resistance of 18 M Ω cm) and concentrations were determined by measuring absorbance at 260 nm. The double-stranded template was formed by annealing the two strands (heat at 95 °C for 5 min, cool 1 °C/min to 25 °C) in 1 \times SPSC buffer (0.4 M NaCl, 50 mM Na₂HPO₄, pH 7.5). RNA probes were transcribed overnight at 37 °C using an AmpliScribe SP6 high yield transcription kit (Epicentre Biotechnologies) with four unmodified ribonucleotide triphosphates. Probes were purified using an RNeasy mini kit (Qiagen) and concentrations were determined by measuring absorbance at 260 nm.

HCR hairpin design. RNA HCR hairpins are 52-nt long (10-nt toehold, 16-bp stem, 10-nt loop). Hairpin sizing was based on *in vitro* and *in situ* binding studies performed in 40% hybridization buffer²⁶. HCR hairpin sequences were designed by considering a set of target secondary structures involving different subsets of the strands (I, H1, H2, I-H1 and I-H1-H2, each as depicted in **Fig. 1a**). For a given target secondary structure, the ensemble defect represents the average number of incorrectly paired nucleotides at equilibrium, calculated over the ensemble of unspseudoknotted secondary structures^{31,32}. Sequence design was performed by mutating the hairpin sequences so as to reduce the sum of the calculated ensemble defects over the set of target structures (J.N. Zadeh, personal communication). Multiple HCR amplifiers were designed independently and then sequence orthogonality was checked using NUPACK (<http://www.nupack.org/>)³³ to simulate the equilibrium species concentrations and base pairing properties for a test tube³⁴ containing different subsets of strands. This approach was used to check for off-target interactions between each of the five initiators and the other four hairpin sets, as well as between the 10-nt toehold and loop segments of each hairpin set and the 10-nt toehold and loop segments of the other four hairpin sets. The sequences are shown in **Supplementary Notes**.

HCR hairpin synthesis. Each HCR hairpin was synthesized by IDT as two segments with one segment end-labeled with an amine (3'-end for H1 and 5'-end for H2) to permit subsequent coupling to a fluorophore. The strand with a 5' end at the ligation site was ordered with a 5' phosphate to permit ligation. Ligation of the two segments produced the full 52-nt hairpin. The ligation was performed using T4 RNA ligase 2 (New England Biolabs) at 16 °C for a minimum of 8 h. The ligated strands were purified using a 15% denaturing polyacrylamide gel. The bands corresponding to the expected sizes of the ligated products were visualized by UV shadowing and excised from the gel. The RNA strands were then eluted by soaking in 0.3 M NaCl overnight and recovered by ethanol precipitation. The pellet was dried and resuspended in ultrapure water and quantified by measuring absorbance at 260 nm. The dye coupling reaction was performed by mixing an amine-labeled hairpin with an Alexa Fluor succinimidyl ester (Invitrogen) and incubating in the dark for 3 h. Alexa-labeled hairpins were separated from unincorporated dyes by repeating the denaturing PAGE purification described above. To ensure that H1 and H2 form hairpin monomers, the strands were snap-cooled in 1 \times SPSC buffer before use (heat at 95 °C for 90 s, cool to room temperature (~23 °C) on the benchtop for 30 min).

Multiplexed gel electrophoresis. Reactions for **Figure 1b** were performed in 40% hybridization buffer (HB) without blocking agents (40% formamide, 2 \times SSC, 9 mM citric acid (pH 6.0), 0.1% Tween 20) with 0.1 μ g/ μ l of total RNA extracted from zebrafish using TRIzol (Invitrogen). Each of the eight hairpin species (two for each of the four HCR amplifiers) was snap-cooled at 3 μ M in 1 \times SPSC buffer. The RNA initiator for each HCR system was diluted to 0.3 μ M in ultrapure water. Each lane was prepared by mixing 12 μ l of formamide, 6 μ l of 5 \times HB supplements without blocking agents (10 \times SSC, 45 mM citric acid (pH 6.0),

0.5% Tween 20), 1.76 μ l of 1.7 μ g/ μ l extracted zebrafish total RNA and 1 μ l of each of the eight hairpins. When an initiator was absent (lane 1), 2.24 μ l of ultrapure water was added to bring the reaction volume to 30 μ l. For lanes 2 to 5, 1 μ l of 0.3 μ M initiator for one HCR amplifier and 1.24 μ l of ultrapure water were added. The reactions were incubated at 45 °C for 1.5 h. The samples were supplemented with 7.5 μ l of 50% glycerol and loaded into a native 2% agarose gel, prepared with 1 \times lithium boric acid buffer (LB) (Faster Better Media). The gel was run at 150 V for 90 min at room temperature and imaged using an FLA-5100 fluorescent scanner (Fujifilm Life Science). The excitation laser sources and emission filters were as follows: a 473 nm laser and a 530 \pm 10 nm bandpass filter (amplifier HCR3, Alexa 488), a 532 nm laser and a 570 \pm 10 nm bandpass filter (amplifier HCR5, Alexa 546), a 635 nm laser and a 665 longpass filter (amplifier HCR1, Alexa 647) and a 670 nm laser and a 705 nm longpass filter (amplifier HCR4, Alexa 700).

***In situ* hybridization studies.** Procedures for the care and use of zebrafish embryos were approved by the Caltech IACUC. Embryos were fixed and permeabilized using the protocol described in **Supplementary Notes**. For the transgenic samples, GFP⁺ embryos were identified using a Leica MZ16 FA fluorescence stereomicroscope. *In situ* hybridization experiments for **Figures 2–4** were performed using the protocol provided in **Supplementary Notes**. Overnight incubations were performed for 16 h. For **Figure 2a–i**, probe solution was prepared by introducing 6 pmol of each probe (1–3 μ l depending on the stock solution) into 300 μ l of 50% HB at 55 °C. Hairpin solution was prepared by introducing 10 pmol of each hairpin (snap-cooled in 5 μ l) into 300 μ l of 40% HB at 45 °C. For **Figure 2j–m**, experiments were performed using WT embryos. A probe set with three probes (1 pmol of each probe) was used for **Figure 2j–l**; probe sets with 1, 3 or 9 probes (1 pmol of each probe) were used for **Figure 2m**. The standard *in situ* protocol was used for both the (AF + NSA) sample (with probes excluded) and for the AF sample (with probes and hairpins excluded). For the (AF + NSA + NSD) sample, *desm* probes were replaced with *egfp* probes carrying the same initiator sequence as the *desm* probes. For the *ex situ* HCR study of **Figure 2k,l**, snap-cooled hairpins (30 pmol of each hairpin) and probes (1 pmol of each probe) were added to 300 μ l of 40% HB and incubated at 45 °C for 16 h while the embryos were incubated without probes in 50% HB at 55 °C. For consistency, these embryos were subjected to the standard probe washes and the standard amplification protocol (substituting the pre-assembled polymer solution for the hairpin solution). Experiments for **Figures 3 and 4** were performed with *Tg(flk1:egfp)* embryos using probe and hairpin solutions prepared following the protocol in **Supplementary Notes**.

Vibratome sectioning. After completion of the standard *in situ* protocol (**Supplementary Notes**), embryos were post-fixed with 4% paraformaldehyde at room temperature for 20 min. Fixation was stopped by washing the embryos three times with 1 \times PBST. Embryos were then embedded in 4% low-melting agarose (Cambrex) in 1 \times PBST and sectioned into 200 μ m slices with a Vibratome Series 1000 tissue sectioning system (Vibratome).

Confocal microscopy. A chamber for mounting the embryo was made by aligning 2 stacks of Scotch tape (eight pieces per stack) 1 cm apart on a 25 mm \times 75 mm glass slide (VWR). Approximately 200 μ l of 3% methyl cellulose mounting medium was added between the tape stacks on the slide and embryos were placed on the medium oriented for lateral imaging. A 22 mm \times 22 mm no. 1 coverslip (VWR) was placed on top of the stacks to close the chamber. The sectioned sample of **Figure 3d** was mounted using a SlowFade Gold antifade reagent (Molecular Probes). A Zeiss 510 upright confocal microscope with an LD LCI Plan-Apochromat 25 \times /0.8 Imm Corr DIC objective was used to acquire the images for **Figure 2**. The excitation laser sources and emissions filters were: 488 nm Ar laser excitation source and a 520 \pm 10 nm bandpass filter (gray; autofluorescence), 633 nm HeNe laser and a 650 nm long pass filter (green; Alexa 647). A Leica TCS SP5 inverted confocal microscope with an HCX PL APO 20 \times /0.7 Imm objective was used to acquire the five-color images of **Figure 3b,d**. Excitation laser sources and tuned emissions bandpass filters were as follows: 488 nm/500–540 nm (Alexa 488), 514 nm/550–565 nm (Alexa 514), 543 nm/550–605 nm (Alexa 546), 594 nm/605–640 nm (Alexa 594), 633 nm/655–720 nm (Alexa 647). Cluster analysis (Leica) was performed

to enhance dye separation. A Zeiss 510 META NLO inverted confocal microscope with an LD C-Apochromat 40×/1.1 W Corr objective was used to acquire the images for **Figure 4**. Excitation laser sources and emission filters were: 488 nm/tunable 500–522 nm (Alexa 488), 514 nm/tunable 543–586 nm (Alexa514), 561 nm/575–630 nm (Alexa 594), 633 nm/650–710 nm (Alexa 647). For the images of **Figure 4**, image registration (rigid body translation and rotation) was performed to correct for possible misalignment between the two channels (TurboReg plugin for ImageJ). All images are presented without background subtraction.

31. Dirks, R.M., Lin, M., Winfree, E. & Pierce, N.A. Paradigms for computational nucleic acid design. *Nucleic Acids Res.* **32**, 1392–1403 (2004).
32. Zadeh, J.N., Wolfe, B.R. & Pierce, N.A. Nucleic acid sequence design via efficient ensemble defect optimization. *J. Compu. Chem.* published online, doi:10.1002/jcc.2163 (17 August 2010).
33. Zadeh, J.N. *et al.* NUPACK: analysis and design of nucleic acid systems. *J. Compu. Chem.* published online, doi:10.1002/jcc.21596 (19 July 2010).
34. Dirks, R.M., Bois, J.S., Schaeffer, J.M., Winfree, E. & Pierce, N.A. Thermodynamic analysis of interacting nucleic acid strands. *SIAM Rev.* **49**, 65–88 (2007).



Cultured cambial meristematic cells as a source of plant natural products

Eun-Kyong Lee^{1,5}, Young-Woo Jin^{1,5}, Joong Hyun Park¹, Young Mi Yoo¹, Sun Mi Hong¹, Rabia Amir², Zejun Yan², Eunjung Kwon^{2,3}, Alistair Elfick³, Simon Tomlinson⁴, Florian Halbritter⁴, Thomas Waibel², Byung-Wook Yun² & Gary J Loake²

A plethora of important, chemically diverse natural products are derived from plants¹. In principle, plant cell culture offers an attractive option for producing many of these compounds^{2,3}. However, it is often not commercially viable because of difficulties associated with culturing dedifferentiated plant cells (DDCs) on an industrial scale³. To bypass the dedifferentiation step, we isolated and cultured innately undifferentiated cambial meristematic cells (CMCs). Using a combination of deep sequencing technologies, we identified marker genes and transcriptional programs consistent with a stem cell identity. This notion was further supported by the morphology of CMCs, their hypersensitivity to γ -irradiation and radiomimetic drugs and their ability to differentiate at high frequency. Suspension culture of CMCs derived from *Taxus cuspidata*, the source of the key anticancer drug, paclitaxel (Taxol)^{2,3}, circumvented obstacles routinely associated with the commercial growth of DDCs. These cells may provide a cost-effective and environmentally friendly platform for sustainable production of a variety of important plant natural products.

Only plant stem cells, embedded in meristems located at the tips of shoots and roots or contained inside the vascular system, can divide and give rise to cells that ultimately undergo differentiation while simultaneously giving rise to new stem cells⁴. These cells can be considered immortal due to their ability to theoretically divide an unlimited number of times. Consequently, since the beginnings of tissue culture in the 1940s, cell suspension cultures have been routinely generated through what was believed to be a dedifferentiation process⁵. Recent evidence suggests this mechanism might not entail a simple reverse reprogramming⁶. Regardless of the mechanism involved, this process results in mitotic reactivation of specialized cell types within a given organ, generating a multicellular mixture of proliferating cells⁷. Suspension cultures derived from such cellular assortments often exhibit poor growth properties with low and inconsistent yields of natural products³, owing to deleterious genetic and epigenetic changes that occur during this process^{7,8}.

To circumvent this so-called dedifferentiation procedure, we developed an innately undifferentiated cell line derived from cambium cells, which function as vascular stem cells⁹. Also, paclitaxel biosynthesis in *T. cuspidata* is most conspicuous within the region containing these CMCs¹⁰. A recently developed twig was collected from a wild yew, *T. cuspidata* (Fig. 1a). We gently peeled tissue that contained cambium, phloem, cortex and epidermis from the xylem (Fig. 1b and Supplementary Fig. 1a–c) and confirmed the absence of xylem cells by staining with phloroglucinol-HCl, which detects lignin deposition (Supplementary Fig. 2a–f). After this tissue was cultured on solid isolation medium for 30 d (Fig. 1c), actively proliferating cambium cells could be gently separated from the DDCs derived from phloem, cortex and epidermis (Fig. 1c–e and Supplementary Fig. 3a–e). This mass of proliferating cells was distinct from DDCs derived from a needle or embryo (Fig. 1f,g), and the morphology of these CMCs differed from adjacent cells (Fig. 1h and Supplementary Fig. 3b–e). We also used this technology to produce such cells from a variety of plant species, including ginseng (*Panax ginseng*), ginkgo (*Ginkgo biloba*) and tomato (*Solanum lycopersicon*). This suggests that the procedure has broad utility (Supplementary Fig. 4a–f).

Microscopic analysis of a suspension culture of *T. cuspidata* cells revealed the presence of small, abundant vacuoles within the cultured cells. This characteristic feature of CMCs¹¹ enables them to withstand the pressure generated by the expanding secondary xylem¹². In contrast, dedifferentiated *T. cuspidata* cells derived from needles or embryos possessed only one large vacuole, typical of such plant cells (Fig. 1i,j). The ability to differentiate into either a tracheary element, the main conductive cell of the xylem, or a phloem element is a defining trait of CMCs^{13,14}. These cultured cells could be conditionally differentiated into a tracheary element at high frequency. In contrast, no tracheary elements were formed from *T. cuspidata* DDCs (Fig. 1k,l). Both animal and plant stem cells are particularly sensitive to cell death triggered by ionizing radiation, to safeguard genome integrity in populations of such cells¹⁵. In a similar fashion, these cultured cells are hypersensitive to γ -irradiation (Fig. 1m) and display increased cell death in response to the radiomimetic drug zeocin¹⁵ (Fig. 1n). In aggregate, our findings, based on a variety of

¹Unhwa Corp., Wooah-Dong, Dukjin-gu, Jeonju, South Korea. ²Institute of Molecular Plant Sciences, School of Biological Sciences, University of Edinburgh, King's Buildings, Edinburgh, UK. ³School of Engineering, University of Edinburgh, King's Buildings, Edinburgh, UK. ⁴Institute for Stem Cell Research, School of Biological Sciences, University of Edinburgh, King's Buildings, Edinburgh, UK. ⁵These authors contributed equally to this work. Correspondence should be addressed to G.J.L. (gloake@ed.ac.uk).

Received 24 August; accepted 27 September; published online 24 October 2010; doi:10.1038/nbt.1693

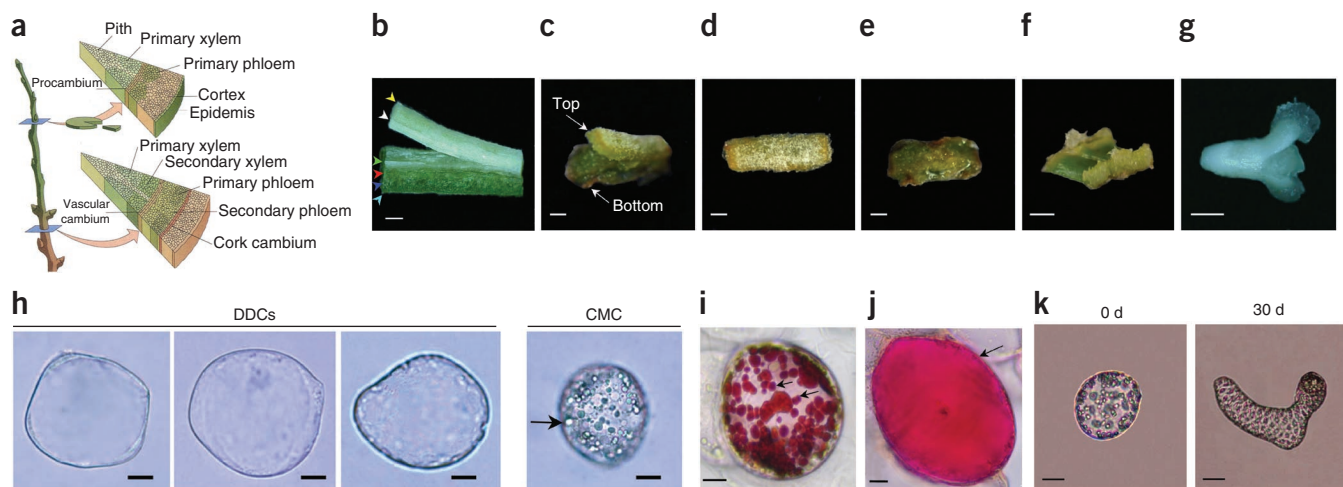


Figure 1 Isolation and culture of *T. cuspidata* CMCs. (a) Schematic cross-section illustrating the location of cambium cells within a typical twig. Reproduced with permission from reference 12. (b) Preparation of *T. cuspidata* explant by peeling off cambium, phloem, cortex and epidermal cells from the xylem. Cell types are indicated by the following colored arrows: yellow, pith; white, xylem; green, cambium; red, phloem; blue, cortex; and turquoise, epidermis. Scale bar, 0.5 mm. (c) Natural split of CMCs from DDCs induced from phloem, cortex and epidermal cells. The top layer is composed of CMCs whereas the bottom layer consists of DDCs. Scale bar, 1 mm. (d) CMCs proliferated from the cambium. Scale bar, 1 mm. (e) DDCs induced from the tissue containing phloem, cortex and epidermal cells. Scale bar, 1 mm. (f) DDCs induced from the cut edge of a needle explant. Scale bar, 0.5 mm. (g) DDCs induced from the cut edge of an embryo explant. Scale bar, 0.5 mm. (h) Micrographs of DDCs and a CMC. CMCs are significantly smaller and possess characteristic numerous, small vacuole-like structures. The black arrow indicates a vacuole-like structure. Scale bars, 20 μ m. (i) Single CMC stained with neutral red, which marks the presence of vacuoles. Two of many stained vacuoles are denoted by black arrows. Scale bar, 10 μ m. (j) Needle-derived DDC stained with neutral red. The single large vacuole present in this cell is marked by a black arrow. Scale bar, 10 μ m. (k) Conditional differentiation of *T. cuspidata* CMCs to tracheary elements, at the times indicated, after addition of differentiation media. Scale bar, 25 μ m. (l) Time-course of differentiation of different *T. cuspidata* cell lines over time into tracheary elements. (m) Quantification of cell death in *T. cuspidata* cells after exposure to increasing levels of ionizing radiation. (n) Levels of cell death in *T. cuspidata* cells after exposure to the radiomimetic drug, Zeocin (phleomycin). Experiments were repeated at least twice with similar results. Data points represent the mean of three samples \pm s.d.

approaches, are consistent with the notion that these cultured cells exhibit stem cell-like properties, consistent with a CMC identity.

We used a combination of deep sequencing technologies to compare the molecular signatures of these cells and those of typical DDCs. First, we used an approach based on massively parallel pyrosequencing¹⁶ to profile the *T. cuspidata* transcriptome. A total of 860,800 reads of average length 351 bp generated 301 MB of sequence (Supplementary Fig. 5a and Supplementary Tables 1–3). From these sequence data, we assembled 36,906 contigs *de novo* (average length, 700 bp; maximum length, 10,355 bp), with 8,865 contigs > 1 kb (Supplementary Fig. 5b and Supplementary Tables 4–6). We subjected contigs from the *T. cuspidata* transcriptome (Supplementary Data Set 1) to BLAST searches and 62% were assigned a putative function (Supplementary Data Set 2). This data set should provide an important resource because there is currently no large-scale sequence information derived from this division of the plant kingdom. The determination of the *T. cuspidata* transcriptome enabled us to use digital gene expression tag profiling¹⁶ to compare gene expression in prospective CMCs with gene expression in DDCs (Supplementary Data Set 3) in the absence of exogenous chemical elicitors that can induce paclitaxel biosynthesis.

Digital gene expression tag profiling analysis established that 563 genes were differentially expressed in CMCs, with 296 upregulated and 267 downregulated (Fig. 2a, Supplementary Figs. 6 and 7 and Data Set 4). A subset of these genes were validated by RT-PCR

(Supplementary Fig. 8). *Phloem intercalated with xylem* (*PXY*) encodes a leucine-rich repeat (LRR) receptor-like kinase (RLK), which is conspicuously expressed in CMCs. *PXY* is a member of a small series of closely related LRR RLKs, mutations that impact CMC function¹⁷. *T. cuspidata* contig 01805 exhibits high similarity to *PXY* (Supplementary Fig. 9a) and is differentially expressed in our prospective CMC suspension cells (Supplementary Data Set 4). Analysis by qRT-PCR established that expression of contig 01805 is upregulated ninefold in these cells relative to DDCs (Fig. 2b).

Wooden Leg (*WOL*) encodes a two-component histidine kinase which is a member of a small gene family in *Arabidopsis*¹⁸. *WOL*-like proteins are unique in having two putative receiver or D-domains and mutations in *WOL* affect vascular morphogenesis¹⁸. *WOL* is expressed in the cambium¹⁸ and *WOL*-like genes are expressed in the cambial zone of the silver birch (*Betula pendula*) and poplar (*Populus trichocarpa*)¹⁹. *T. cuspidata* contig 10710 exhibits high similarity to *WOL* and its related genes (Supplementary Fig. 9b). Gene expression analysis established that this gene is upregulated 12-fold in CMCs relative to DDCs (Fig. 2b).

To assess the molecular composition and the relative expression of genes indicative of given biological pathways in our prospective CMCs, we performed enrichment analysis of Gene Ontology (GO) terms within our data set. This approach established that both stress and biotic defense response genes were prominently over-represented

Figure 2 Characterization of CMCs from *T. cuspidata*, including transcriptome profiling using digital gene expression tags. **(a)** Scatter plot showing differentially expressed genes (DEGs) (blue and red) in CMCs and non-DEGs (black). The deployment of further filtering approaches identified more robust DEGs (red), whereas other DEGs (blue) were filtered out. $FDR \leq 0.05$; $n = 1,229$. **(b)** Analysis of the expression of contig 01805 and contig 10710. **(c)** Relative percentage of GO groups within CMC DEGs. **(d)** Growth of CMCs and DDCs derived from needles or embryos on solid growth media from an initial 3 g f.c.w. 95% confidence limits are too small to be visible on this scale. **(e)** Bar graph reporting the extent of cell aggregation in DDC and CMC suspension cultures. **(f)** Paclitaxel production by 3-month-old DDCs and CMCs 10 d after elicitation, following batch culture in a flask format. Error bars represent 95% confidence limits. These experiments were repeated three times with similar results.

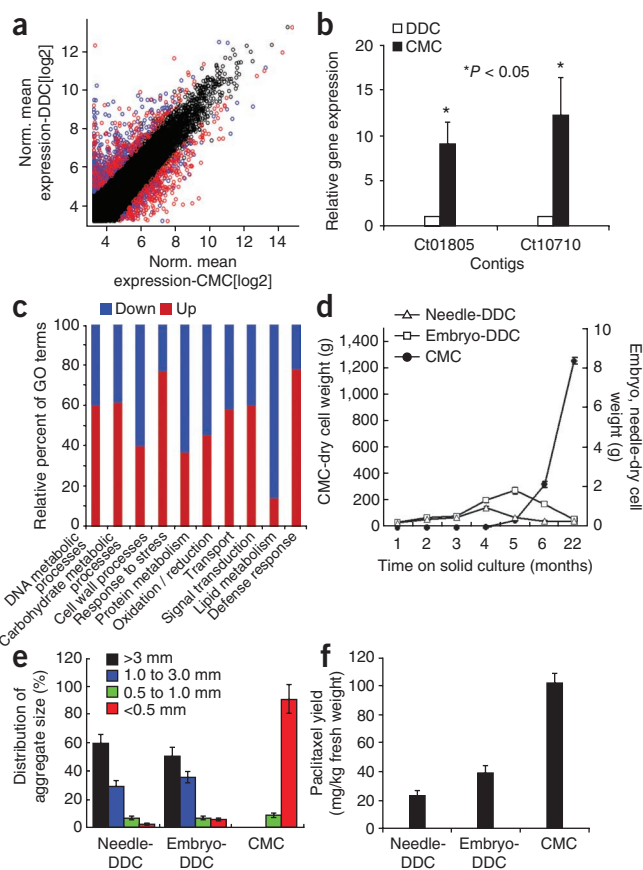
(Fig. 2c). Stem cells exhibit a low threshold for auto-execution through apoptosis but express robust defenses against environmental stresses²⁰. Collectively, our Digital gene expression tag profiling data are therefore consistent with a CMC identity for these cultured cells.

We used a solid growth media format and representative cell lines to compare the growth properties of these CMCs with DDCs derived from the same wild tree. At 22 months after inoculation, the *T. cuspidata* needle- and embryo-derived DDCs produced a total dry cell weight (d.c.w.) of 0.32 g and 0.41 g, respectively, when grown on solid media with subculturing every 2 weeks (Fig. 2d). In contrast, the d.c.w. generated from CMCs was 1,250 g, an increase of $4.0 \times 10^5\%$ and $3.0 \times 10^5\%$, respectively. Moreover, these cells were still growing rapidly following 22 months of culture, whereas DDCs possessed conspicuous necrotic patches (Supplementary Fig. 10) and displayed signs of a rapid decrease in their viable cell biomass.

Pronounced cell aggregation is a typical feature of suspension cultures comprised of DDCs. This can lead to differences in local environments between cells significantly reducing growth rate and natural product biosynthesis³. In representative suspension cultures of DDCs derived from either *T. cuspidata* needles or embryos, the proportion of cell aggregates with a diameter <0.5 mm was only 2% or 5%, respectively ($n = 150$ cells). Conversely, in representative CMCs, 93% of cell aggregates had a diameter <0.5 mm, with many cells present as singletons or components of aggregates comprised of only 2–3 cells ($n = 150$ cells) (Fig. 2e and Supplementary Fig. 11).

We next determined the magnitude of paclitaxel biosynthesis in this novel cell line during batch culture in a 125 ml Erlenmeyer flask. At 14 d after inoculation of flask cultures with cells, cells were transferred to production medium containing the elicitors methyl-jasmonate² and chitosan, together with a precursor phenylalanine, to induce paclitaxel biosynthesis. Levels of paclitaxel were measured 10 d later by high-performance liquid chromatography (HPLC). The amount of paclitaxel produced, 102 mg/kg fresh cell weight (f.c.w.), was conspicuously greater than that generated by either needle or embryo-derived DDCs at a f.c.w. value of 23 mg/kg or 39 mg/kg, respectively (Fig. 2f). Measurements of this natural product were confirmed by liquid chromatography mass spectrometry (LC-MS) (Supplementary Fig. 12a–d). Further, genes encoding key enzymes integral to the biosynthesis of paclitaxel^{2,3} were induced more strongly in CMCs than in DDCs (Supplementary Fig. 13).

To establish whether these cells exhibit superior growth properties on a larger scale, we first investigated their performance in a 10 liter stirred tank bioreactor. In this environment, shear stress can limit growth, and the problem is often intensified by cell aggregation²¹. The CMCs in this bioreactor grew significantly faster than DDCs (Fig. 3a). Further, in response to shear stress, the survival rate of CMCs was strikingly higher than for DDCs, which by the



end of the culture period had largely turned necrotic and had stopped growing (Supplementary Fig. 14a–c).

Next, we explored the performance of these cells in a 3 liter air-lift bioreactor. Large aggregates of DDCs formed in this bioreactor, leading to reduced cell mixing and circulation, which subsequently resulted in cell adherence to the bioreactor wall. Furthermore, many of these adhered cellular aggregates developed necrotic patches. After 4 months of culture, the growth of DDCs from either needle or embryo, expressed as dry cell weight (d.c.w.), were 3.33 g and 5.08 g, respectively. In contrast, the CMC line had generated a d.c.w. of 3,819.44 g, an increase of 114,000% and 75,000%, respectively (Fig. 3b). We also analyzed the growth of these cell lines in a 20 liter air-lift bioreactor, routinely used as a pilot for subsequent large-scale production. DDCs did not grow in this size bioreactor under the conditions tested and rapidly became necrotic. Conversely, CMCs always grew rapidly, increasing their d.c.w. from 3.65 g/l to 12.85 g/l within 14 d (Fig. 3c). Their relative tolerance of shear stress can likely be attributed to their small and abundant vacuoles, reduced aggregation and thin cell walls²¹.

We attempted to improve the performance of needle- and embryo-derived DDCs by specifically selecting the more rapidly growing cells at each subculture on solid media for a period of 1.8 years. This process improved the growth of needle-, but not embryo-derived, DDCs in a 3 liter air-lift bioreactor. However, the performance of CMCs was still strikingly superior to that of these extensively selected cells with respect to specific growth rate (μ), doubling time (T_d) and growth index (GI) (Supplementary Fig. 15 and Supplementary Table 7). A key trait for the exploitation of plant cells on an industrial scale is the stability of their growth in suspension culture³. We therefore monitored the growth stability of these cells compared to selected

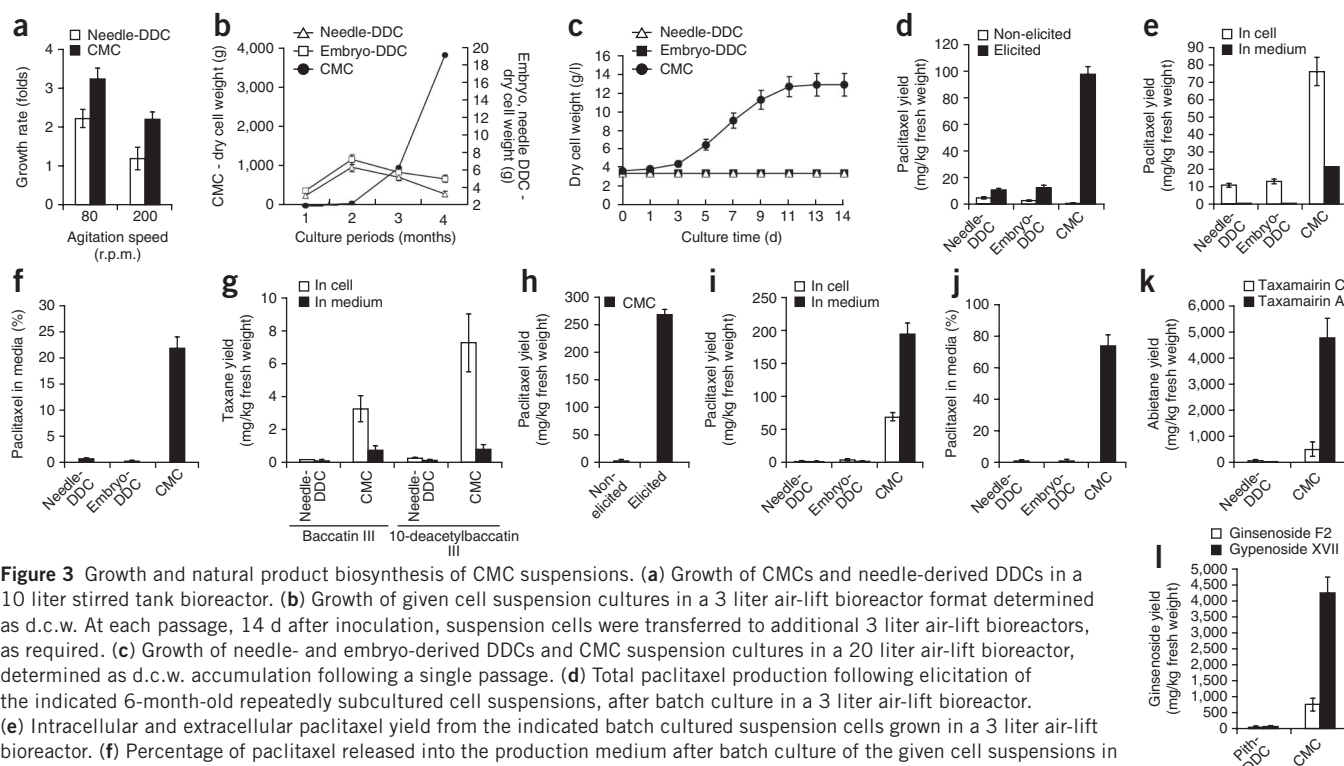


Figure 3 Growth and natural product biosynthesis of CMC suspensions. **(a)** Growth of CMCs and needle-derived DDCs in a 10 liter stirred tank bioreactor. **(b)** Growth of given cell suspension cultures in a 3 liter air-lift bioreactor format determined as d.c.w. At each passage, 14 d after inoculation, suspension cells were transferred to additional 3 liter air-lift bioreactors, as required. **(c)** Growth of needle- and embryo-derived DDCs and CMC suspension cultures in a 20 liter air-lift bioreactor, determined as d.c.w. accumulation following a single passage. **(d)** Total paclitaxel production following elicitation of the indicated 6-month-old repeatedly subcultured cell suspensions, after batch culture in a 3 liter air-lift bioreactor. **(e)** Intracellular and extracellular paclitaxel yield from the indicated batch cultured suspension cells grown in a 3 liter air-lift bioreactor. **(f)** Percentage of paclitaxel released into the production medium after batch culture of the given cell suspensions in a 3 liter air-lift bioreactor. **(g)** Synthesis of baccatin III and 10-deacetyl/baccatin III in CMCs relative to needle-derived DDCs. **(h)** Magnitude of paclitaxel biosynthesis following elicitation of 28-month-old CMCs in a 20 liter air-lift bioreactor. Needle- and embryo-derived DDC suspensions did not routinely grow in this size bioreactor. **(i)** Intracellular and extracellular paclitaxel yield after 45 d of perfusion of cultured needle- and embryo-derived DDCs and CMCs in a 3 liter air-lift bioreactor. **(j)** Percentage of paclitaxel released into the production medium after perfusion culture of the given cell suspensions as indicated in i. **(k)** Synthesis of taxamairin A and C in CMCs and needle-derived DDCs after batch culture in a 3 liter air-lift bioreactor. **(l)** Synthesis of ginsenosides in *P. ginseng* CMC and pith-derived DDC suspension cells after batch culture in a 3 liter air-lift bioreactor. Error bars represent 95% confidence limits. These experiments were repeated twice with similar results.

DDCs derived from needles. The CMCs exhibited a relatively constant growth rate over time. In contrast, we observed striking fluctuations in growth rates during the culture of DDCs (**Supplementary Fig. 16**). Finally, we determined the growth of CMCs within a 3 ton bioreactor. These cells were again successfully cultured with high performance (**Supplementary Fig. 17a and b**), establishing their utility for growth on an industrial scale.

We determined the level of paclitaxel production of these different *T. cuspidata* cell suspensions in both 3 liter and 20 liter air-lift bioreactors. At 10 d after elicitation, CMCs again synthesized strikingly more paclitaxel than either of the DDC lines in a 3 liter air-lift bioreactor. Further, elicitation induced a 220% (11 mg/kg) and 433% (13 mg/kg) increase in paclitaxel production within needle- and embryo-derived DDCs respectively, whereas the induction was 14,000% (98 mg/kg) with CMCs (**Fig. 3d**). CMCs secreted $2.7 \times 10^4\%$ and $7.2 \times 10^4\%$ more paclitaxel into the culture medium than the low levels secreted by either needle- or embryo-derived DDCs, respectively (**Fig. 3e,f**). The amount of paclitaxel secreted to the medium during culture varies significantly both between *Taxus* species and in response to different culture conditions²². Our DDCs secreted less paclitaxel than might be expected. Nevertheless, *T. cuspidata* CMCs secreted a strikingly greater amount of paclitaxel into the medium under these culture conditions than the associated DDCs. Moreover, these cells also synthesized strikingly more of the related taxanes baccatin III and 10-deacetyl/baccatin III^{2,3} (**Fig. 3g**). No paclitaxel production was detected by either DDC line in a 20 liter air-lift bioreactor. In contrast, CMCs synthesized 268 mg/kg and were again highly responsive to elicitation (**Fig. 3h**). Previously

reported values for *T. cuspidata* paclitaxel production range from 20–84 mg/kg f.c.w.^{23,24}. However, these data, including the maximum value, were obtained from flask cultures, whereas our data suggest that DDCs have improved function relative to their performance on a larger scale. Our findings imply that CMCs synthesize strikingly more paclitaxel and are significantly more responsive to elicitation when batch cultured in either 3 liter or 20 liter air-lift bioreactors compared to typical *T. cuspidata* suspension cells.

Perfusion culture promotes the secretion of secondary metabolites into the culture medium, aiding both purification and natural product biosynthesis²². We therefore compared the magnitude of paclitaxel secretion after perfusion culture. Following 45 d of perfusion culture, needle- and embryo-derived DDCs were largely necrotic; however, CMCs produced a combined total of 264 mg of paclitaxel per kg of cells and 74% of this was secreted directly into the medium (**Fig. 3i,j**). Perfusion culture of these cells therefore both promotes paclitaxel biosynthesis and increases the proportion of this secondary product that is secreted into the medium, facilitating its cost-effective purification. The future deployment of metabolic engineering approaches and higher yielding *Taxus* species may further enhance paclitaxel biosynthesis in these cells^{2,3}.

We also monitored these *T. cuspidata* suspension cultures for the production of the abietane tricyclic diterpenoid derivatives, taxamairin A and taxamairin C, which have also been shown to possess antitumor activities²⁵. Elicitation of these cells within a 3 liter air-lift bioreactor induced increases in both taxamairin C and especially taxamairin A to 520.8 and 4,982.5 mg/kg f.c.w., respectively, in CMCs. These values were far greater than those determined in

DDCs (Fig. 3k). Suspension cultures of *T. cuspidata* have previously been reported to produce 0.92 and 26.08 mg/kg f.c.w. of taxamairin C and taxamairin A, respectively²⁶. Our data imply that CMCs might provide a considerably better source of these abietanes than DDCs. To establish whether CMCs derived from other plant species also exhibit superior properties with respect to the biosynthesis of commercially relevant natural products, we determined the synthesis of ginsenosides, a class of triterpenoid saponins derived exclusively from the plant genus *Panax*. Ginsenosides have been reported to show multiple bioactivities including neuroprotection, antioxidative effects and the modulation of angiogenesis²⁷. Following elicitation of tap root-derived *P. ginseng* suspension cells, cultured using a 3 liter air-lift bioreactor, ginsenoside F2 and gypenoside XVII accumulated to strikingly greater levels in *P. ginseng* CMCs relative to DDCs. Ginsenoside F2 and gypenoside XVII accrued to 791 and 4,425 mg/kg f.c.w., respectively (Fig. 3l). Previously, ginsenoside F2 has been reported to reach 33.3 mg/kg f.c.w.²⁸ and gypenoside XVII 183.3 mg/kg f.c.w.²⁹ in ginseng roots. Thus, CMCs synthesize 23.8- and 24.1-fold more ginsenoside F2 and gypenoside XVII, respectively, than previously described sources. Therefore, CMCs may also be used for the production of ginsenosides.

Numerous medicines, perfumes, pigments, antimicrobials and insecticides are derived from plant natural products^{1–3,30}. Cultured cambial meristematic cells may provide a cost-effective, environmentally friendly and sustainable source of paclitaxel and potentially other important natural products. Unlike plant cultivation, this approach is not subject to the unpredictability caused by variation in climatic conditions or political instability in certain parts of the world. Furthermore, CMCs from reference species may also provide an important biological tool to explore plant stem cell function.

METHODS

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

Accession codes. Sequence Read Archive: ERP000352.

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

ACKNOWLEDGMENTS

T.W. was awarded a BBSRC CASE PhD studentship. This project was funded in part by a grant from the Korea Institute for Advancement of Technology (KIAT) (R & D project number: 10030175), the Ministry of Knowledge Economy (MKE), Republic of Korea to E.-K.L., J.H.P., S.M.H. and G.J.L. R.A. was supported by a scholarship from HEC Pakistan. E.K. was supported by a studentship from the Engineering and Physical Sciences Research Council. We acknowledge the expert technical assistance of A. Montazam and D. Cleven for Roche 454 sequencing and M. Thomson for Illumina Solexa sequencing. Further, S. Bridgett and U. Trivedi provided invaluable input for bioinformatic analysis of the deep sequencing data. All sequencing was undertaken at the GenePool facility, University of Edinburgh.

AUTHOR CONTRIBUTIONS

E.-K.L., Y.-W.J., J.H.P., T.W. and B.-W.Y. performed experiments. R.A., E.K., S.T. and F.H. contributed to bioinformatic and statistical analysis. Z.Y., Y.M.Y. and S.M.H. carried out experiments. A.E. co-supervised E.K. E.-K.L., Y.-W.J. and G.J.L. formulated experiments. G.J.L. and E.-K.L. wrote the paper. All authors discussed results and commented on 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/>.

- Schmidt, B.M., Ribnicky, D.M., Lipsky, P.E. & Raskin, I. Revisiting the ancient concept of botanical therapeutics. *Nat. Chem. Biol.* **3**, 360–366 (2007).
- Croteau, R., Ketchum, R.E.B., Long, R.M., Kaspera, R. & Wildung, M.R. Taxol biosynthesis and molecular genetics. *Phytochem. Rev.* **5**, 75–97 (2006).
- Roberts, S.C. Production and engineering of terpenoids in plant cell culture. *Nat. Chem. Biol.* **3**, 387–395 (2007).
- Laux, T. The stem cell concept in plants: a matter of debate. *Cell* **113**, 281–283 (2003).
- Thorpe, T.A. History of plant tissue culture. *Mol. Biotechnol.* **37**, 169–180 (2007).
- Sugimoto, K., Jiao, Y. & Meyerowitz, E.M. *Arabidopsis* regeneration from multiple tissues occurs via a root development pathway. *Dev. Cell* **18**, 463–471 (2010).
- Grafi, G. et al. Histone methylation controls telomerase-independent telomere lengthening in cells undergoing dedifferentiation. *Dev. Biol.* **306**, 838–846 (2007).
- Baehler, S. et al. Establishment of cell suspension cultures of yew (*Taxus x Media* Rehder) and assessment of their genomic stability. *In Vitro Cell. Dev. Biol. Plant* **41**, 338–343 (2005).
- Ye, Z.-H. Vascular tissue differentiation and pattern formation in plants. *Annu. Rev. Plant Biol.* **53**, 183–202 (2002).
- Strobel, G.A. et al. Taxol formation in yew-*Taxus*. *Plant Sci.* **92**, 1–12 (1993).
- Frankenstein, C., Eckstein, D. & Schmitt, U. The onset of cambium activity - A matter of agreement? *Dendrochronologia* **23**, 57–62 (2005).
- Moore, R., Clark, W.D., Stern, K.R. & Vodopich, D. (eds.) *Botany* (Wm.C. Brown, Dubuque, Iowa, USA; 1995).
- Turner, S., Gallois, P. & Brown, D. Tracheary element differentiation. *Annu. Rev. Plant Biol.* **58**, 407–433 (2007).
- Ito, Y. et al. Dodeca-CLE peptides as suppressors of plant stem cell differentiation. *Science* **313**, 842–845 (2006).
- Fulcher, N. & Sablowski, R. Hypersensitivity to DNA damage in plant stem cell niches. *Proc. Natl. Acad. Sci. USA* **106**, 20984–20988 (2009).
- Shendure, J. & Ji, H. Next-generation DNA sequencing. *Nat. Biotechnol.* **26**, 1135–1145 (2008).
- Fisher, K. & Turner, S. PXY, a receptor-like kinase essential for maintaining polarity during plant vascular-tissue development. *Curr. Biol.* **17**, 1061–1066 (2007).
- Mähönen, A.P. et al. A novel two-component hybrid molecule regulates vascular morphogenesis of the *Arabidopsis* root. *Genes Dev.* **14**, 2938–2943 (2000).
- Nieminen, K. et al. Cytokinin signaling regulates cambial development in poplar. *Proc. Natl. Acad. Sci. USA* **105**, 20032–20037 (2008).
- Rando, T.A. The immortal strand hypothesis: segregation and reconstruction. *Cell* **129**, 1239–1243 (2007).
- Joshi, J.B., Elias, C.B. & Patole, M.S. Role of hydrodynamic shear in the cultivation of animal, plant and microbial cells. *Chem. Eng. J.* **62**, 121–141 (1996).
- Wang, C., Wu, J. & Mei, X. Enhanced taxol production and release in *Taxus chinensis* cell suspension cultures with selected organic solvents and sucrose feeding. *Biotechnol. Prog.* **17**, 89–94 (2001).
- Mirjalili, N. & Linden, J.C. Methyl jasmonate induced production of Taxol in suspension cultures of *Taxus cuspidata*: Ethylene interaction and induction models. *Biotechnol. Prog.* **12**, 110–118 (1996).
- Wu, Z.L., Yuan, Y.-J., Ma, Z.-H. & Hu, Z.D. Kinetics of two-liquid-phase *Taxus cuspidata* cell culture for production of Taxol. *Biochem. Eng. J.* **5**, 137–142 (2000).
- Yang, S.-J., Fang, J.-M. & Cheng, Y.-S. Lignans, flavonoids and phenolic derivatives from *Taxus mairei*. *J. Chinese Chem. Soc.* **46**, 811–818 (1999).
- Bai, J. et al. Production of biologically active taxoids by a callus culture of *Taxus cuspidata*. *J. Nat. Prod.* **67**, 58–63 (2004).
- Leung, K.W. & Wong, A.S.-T. Pharmacology of ginsenosides: a literature review. *Chin. Med.* **5**, 20 (2010).
- Dan, M. et al. Metabolite profiling of *Panax notoginseng* using UPLC-ESI-MS. *Phytochemistry* **69**, 2237–2244 (2008).
- Reynolds, L.B. Effects of harvest date on some chemical and physical characteristics of American ginseng (*Panax quinquefolius* L.). *J. Herbs Spices Med. Plants* **6**, 63–69 (1998).
- Kutchan, T. & Dixon, R.A. Physiology and metabolism: Secondary metabolism: nature's chemical reservoir under deconvolution. *Curr. Opin. Plant Biol.* **8**, 227–229 (2005).



ONLINE METHODS

Collection and sterilization of *T. cuspidata* samples. Twig, needle and seed samples were collected from a wild-grown *T. cuspidata* tree. Twig and needle samples were immediately deposited in 0.56 mM ascorbic acid solution. They were stored at 4 °C for 1 month. Then, they were washed in running tap water for 30 min and surface disinfected with 70% ethanol for 1 min, followed by 1% sodium hypochlorite for 20 min for twigs and 15 min for needles and 0.07% NaOCl for 20 min, and rinsed 5 times with sterilized distilled water (dH₂O). Lastly, they were rinsed once with dH₂O containing 150 mg/l citric acid. Seeds were put into 0.01% NaOCl for 24 h with agitation. They were washed in running tap water for 4 h, surface disinfected with 70% ethanol for 1 min and then placed in a 1% NaOCl for 15 min. Then, they were rinsed 5 times with dH₂O.

Isolation of CMCs. For CMCs, cambium, phloem, cortex and epidermal tissue were peeled off from the xylem and the epidermal tissue side was laid on B5 medium³¹ excluding (NH₄)₂SO₄ with 1 mg/l picloram, 30 g/l sucrose and 4 g/l gelrite. After 4 to 7 d, cell division was evident only in cambium and after 15 d, DDCs started to form from the layer that consisted of phloem, cortex and epidermis by dedifferentiation. At 30 d post-culture, there was a visible split between cambium cells and DDCs of the phloem, cortex and epidermis. This separation was obvious because cambium cells uniformly divided resulting in the formation of a flat plate of cells. In contrast, DDCs derived from phloem, cortex and epidermis proliferated in an irregular form, presumably due to the discrepancy between cell division rates. Following the natural separation of cambium from the other cell types, this cell layer was transferred onto different Petri dishes containing B5 medium excluding (NH₄)₂SO₄ with 1 mg/l picloram, 10 g/l sucrose and 4 g/l gelrite. Initial cell inoculum size was 3.0 g (f.c.w.) and subsequently, CMCs were subcultured onto the fresh medium every 2 weeks. Establishment of *P. ginseng* CMCs was as described above except that the isolation medium contained McCown woody plant medium with 2 mg/l IAA, 30 g/l sucrose, 100 mg/l ascorbic acid, 150 mg/l citric acid and 3 g/l gelrite. For lignin visualization, tissues were stained with phloroglucinol-HCl (0.5% (wt/vol) phloroglucinol in 6 N HCl) for 5 min and then observed under a light microscope.

Establishment of cell suspension cultures and natural product production. Initial suspension cultures were established by inoculating a sample of 2.5 g (f.c.w.) cultured cells derived from either cambium, needles or embryos into 125 ml Erlenmeyer flasks containing 25 ml of B5 medium containing 1 mg/l picloram, and 20 g/l sucrose, excluding (NH₄)₂SO₄. The flasks were agitated at 100 r.p.m. and 21 °C in the dark. Subculturing was undertaken at 2-week intervals.

For culturing the cells in 3 liter and 20 liter air-lift bioreactors, the same medium that was used in the initial suspension culture was applied. Diameter, height and pore size of micro-sparger used in the bioreactor was 2 cm, 0.4 cm, 0.2 μm, respectively. Aeration rate was 0.1–0.2 vol/vol/min (v.v.m.) in 3 liter air-lift bioreactor, and 0.08–0.18 v.v.m. in 20 liter air-lift bioreactor. 3.25 g/l (d.c.w.) of CMCs, 3.3 g/l (d.c.w.) of needle-derived DDCs and 3.1 g/l (d.c.w.) of embryo-derived DDCs were inoculated in 3 liter air-lift bioreactor. 3.65 g/l (d.c.w.) of CMCs, 3.64 g/l (d.c.w.) of needle-derived DDCs and 3.41 g/l (d.c.w.) of embryo-derived DDCs were inoculated in 20 liter air-lift bioreactor. Working volume was 80% of total capacity, which is 2.4 liters in 3 liter air-lift bioreactor and 16 liters in 20 liter air-lift bioreactor. Subculturing of CMCs and DDCs was undertaken every 2 weeks in 3 liter and 20 liter air-lift bioreactor with same initial inoculum size and conditioned medium was recycled with the ratio of 25% of working volume. Growth rate was measured in d.c.w. (g/l) after vacuum filtration and drying of the cells in an oven at 70 °C for 24 h. We call these CMCs Ddobyul, meaning ‘another star’ in Korean.

To test the capacity for production of paclitaxel in the flask and 3 liter and 20 liter air-lift bioreactors, cells at 14 d of culturing were transferred to B5 medium excluding KNO₃, and containing 60 g/l fructose and 2 mg/l 1-naphthalene acetic acid (NAA), and elicitors such as 50 mg/l, chitosan and 100 μM methyl-jasmonate, in addition to 0.1 mM of the precursor, phenylalanine.

Me-JA was dissolved in 90% ethanol, chitosan in glacial acetic acid and phenylalanine in distilled water before dilution to the required concentrations. After 10 d, paclitaxel content was analyzed. Taxane and abietane production was

elicited in a similar fashion. Stress-triggered ginsenoside accrual was mediated by reducing air supply from a constant 0.1 v.v.m. into a 3 liter air-lift bioreactor, for 13 d of culture, to 0.1 v.v.m. for a 30 min period twice per day for 3 d.

Establishment of needle and embryo DDC cultures. DDCs were induced from embryos and needles largely as previously described^{32,33}. For induction of needle-derived DDCs, both ends of the needle were cut in 0.3–0.5 cm (length and width) and were laid on B5 medium containing 1 mg/l picloram, 30 g/l sucrose and 4 g/l gelrite, excluding (NH₄)₂SO₄. After 30 d of culturing, DDCs were induced from the cut edges (Fig. 1f). As culture period continued, DDCs formed over the whole explants. Induced DDCs were transferred to B5 medium containing 1 mg/l picloram, 10 g/l sucrose and 4 g/l gelrite, excluding (NH₄)₂SO₄ for growth. Initial inoculum size was 3.0 g (f.c.w.) and DDCs were subcultured to fresh medium every 2 weeks.

For induction of embryo-derived DDCs, both ends of the zygotic embryo were cut and laid on B5 medium containing 1 mg/l 2,4-D, 30 g/l sucrose and 4 g/l gelrite, excluding (NH₄)₂SO₄. After 23 d of culturing, DDCs were induced from the cut edges (Fig. 1g). As culture period continued, DDCs formed over the whole explant. Induced DDCs were transferred to B5 medium containing 1 mg/l picloram, 10 g/l sucrose and 4 g/l gelrite, excluding (NH₄)₂SO₄ for growth. Initial inoculum size was 3.0 g (f.c.w.) and DDCs were subcultured to fresh medium every 2 weeks.

CMC differentiation. The media used to induce CMC differentiation into tracheary elements (TEs) was B5 medium, excluding (NH₄)₂SO₄ with 10 mg/l NAA, 2 mg/l kinetin, 6 mg/l GA₃ and 60 g/l sucrose. TEs were identified by virtue of their bright fluorescence, due to the presence of lignified secondary cell walls. The extent of TE differentiation was determined as the percentage of TEs per total number of cells. This analysis was undertaken in triplicate and in each case 200 cells were counted.

Response of *T. cuspidata* cell suspensions to γ-irradiation and radiomimetic drugs. CMCs and needle-derived DDCs were obtained from suspension cultures obtained from 20 liter air-lift bioreactors. For gamma-irradiation (Co⁶⁰), cells were irradiated at a dose rate of 0.92 Gy/min for 0–400 Gy, which has been modified from a method described previously³⁴. Then, cells were suspension cultured for 24 h in 100 ml flasks at 21 °C, 100 r.p.m. in the dark (volume of cells to liters of medium was 1:10). Suspension cells were treated with Zeocin (200 μg/ml, Invitrogen) at 7 d after subculture, essentially as described previously³⁴. The treated suspension cell culture was incubated in the dark for 24 or 48 h. For cell death determination, cells were treated with 2% Evan's blue for 3 min and washed with sterile water several times, then transferred to a microscope slide covered with a thin cover slip. For each sample, cell death was determined 5 times independently and the average cell death rate was measured by excluding the maximum and minimum number of cell counts. All experiments were undertaken in triplicate.

Determination of *T. cuspidata* transcriptome. RNA was isolated using a Qiagen plant RNA kit following the manufacturer's instructions. cDNA was synthesized by employing a SMART procedure to enrich for full-length sequences³⁵. The resulting cDNA was normalized using kamchatka crab duplex-specific nuclease³⁶, to aid the discovery of rare transcripts. cDNA was sheared using Covaris instruments settings: target size 500 bp, duty cycle 5%, intensity 3, cycles/burst 200 and time 90 s. Library preparation was undertaken using a Roche GS FLX library kit. The concentration and quality of the synthesized library was analyzed using a Agilent bio-analyzer. Titration emulsion PCR using a GS FLX emPCR kit was undertaken to determine the optimum number of beads to load for large-scale sequencing. A Beckman/Coulter Multisizer 3 bead counter was employed to determine the concentration of beads. Two million beads were loaded onto a GS FLX pico titer plate using a Roche 70 × 75 kit.

The sequencing reagents used and washes undertaken followed protocols from the manufacturer. The *T. cuspidata* transcriptome was determined in the GenePool genomics facility at the University of Edinburgh using a Roche 454 GS FLX instrument in titanium mode, which uses massively parallel pyrosequencing technology^{37,38}. A total of 860,800 reads were achieved of 351 bp average length, which generated 301 MB of sequence. These data were

assembled into isotigs by employing Newbler 2.3. BLAST was used to search for similar sequences within available sequence databases. Annot8r was employed to predict GO terms for each isotig³⁹.

Digital gene expression tag profiling. The analysis of global gene expression in *T. cuspidata* cell suspension cultures was carried out by digital gene expression tag profiling, using an improved method based on previously described technology⁴⁰. Potentially contaminating DNA was removed from RNA samples using Ambion turbo DNase treatment. NlaIII library preparation was accomplished by following the manufacturer's standard protocol. Fifteen PCR cycles were used for amplification. We used 1–10 µg of a given library for sequencing from each sample. Sequencing was carried out in the GenePool genomics facility at the University of Edinburgh using a Genome Analyser (GA) II_x Illumina Solexa sequencing machine. Three replicates each of both cell lines (CMC/DDC) were sequenced following the manufacturer's protocol. Subsequently, reads were aligned to the *T. cuspidata* reference genome using MAQ v. 6.0.8. and uniquely aligned reads to the previously assembled *T. cuspidata* transcriptome were counted.

Statistical analysis. Statistical analysis was performed in R using the edgeR Bioconductor library^{40,41}. We sought to reduce problems created by varying library sizes and noise for genes not highly expressed, by stabilizing read counts through adding a small constant. Therefore we first rescaled the read counts in each library by dividing by the sum of all read counts in the upper quartile of expression values⁴² and afterwards added a constant factor C (C = 10) to each count. This transformation alters the signal in such a way that differences between groups for contigs with low expression are less likely to be considered differentially expressed, while leaving high transcript counts largely untouched. Briefly, edgeR uses an overdispersed Poisson distribution to model read count data, where the degree of overdispersion is moderated using an empirical Bayes procedure. Differential expression is assessed using a modified version of Fisher's exact test. We ran edgeR according to the steps outlined in the library's tutorial (using parameter settings prior. $n = 10$, grid.length = 500). P-values were adjusted for the false discovery rate and we deemed a threshold of false discovery rate (FDR) ≤ 0.05 to be appropriate to detect differentially expressed contigs ($n = 1,229$).

In the latter analysis, we decided to first focus on only those differentially expressed contigs, that showed a considerably large change (that is, the minimum difference between any replicates in both groups, DDC and CMC, was at least ten transcripts per million) and for which the direction of change was consistent between all replicates (that is, all replicates are either higher or all replicates are lower in one group than in the other). We considered these filtered contigs ($n = 563$) the most interesting candidates for immediate study and held out the rest for further follow-up studies.

Gene expression analysis. The determination of gene expression levels were carried out by either RT-PCR or qRT-PCR as previously described⁴³. The primer sequences employed are listed below:

Primer sequences for qRT-PCR.

Ct01805-F: CTTGGCAAGGATCCAGTTTAG

Ct01805-R: AGACCAAGCCCAGGGTCTTC

Ct10710-F: TTCTTCGGCTGTGTCAGTGATG

Ct10710-R: CCGATAGAAGCTTGACAGAA

Primer sequences for RT-PCR.

Ct27072-F: CACTTGGAGTTCGTCGTTGA

Ct27072-R: CACTGTGCACACTACCCAAA

Ct36802-F: GAGCCGTTGCATGGTACACT

Ct36802-R: TAACCGTGGTGTCAAATCA

Ct18649-F: CCTGACAACAGCGTCTCTGA

Ct18649-R: AAACCACAGTACCCACAGC

Ct33753-F: GTTAGACCCTTACCCTGCA

Ct33753-R: CTGCAAAGCTGAGAGTGGAAATG

Ct30863-F: GCAACGTCGAAACGCAGTA

Ct30863-R: AGAGTTGCGAACAGCAAAGG

Ct34959-F: ACTCGATAGACCGACAAGG

Ct34959-R: CAGCTGATCGTCCAGCTATG

Ct01720-F: CTCCTCTCAACGAGGAAAA

Ct01720-R: GTTTTCCCAGAAGGGAATC

Ct09814-F: TTGAGGCATGTGGGTTTAA

Ct09814-R: TGTCATCTGTTGCATTGGA

Ct07968-F: CGACAACATTCTTGCATTGA

Ct07968-R: AACCGTTGCAGGGAACCTAC

Ct03409-F: ATGTTCCAAAAATGGGAGGA

Ct03409-R: GCTTGGAAAAGACCTGAAGGA

Ct04884-F: AGTGAATGTAAGCCCATGA

Ct04884-R: TTTGGCATCTTCTTGATGA

Ct07286-F: GTCCATCCATTGTCCATAGAAA

Ct07286-R: TGGCAACATTGGTAAAGATATTCA

Perfusion culture. Perfusion culture was initiated in a similar fashion to that described for the bioreactors. On day 14, cultures were elicited with 50 mg/l chitosan, 0.1 mM phenylalanine and 100 µM methyl jasmonate. After elicitation, the spent medium was removed aseptically and replaced with an equal volume of fresh B5 medium excluding KNO₃ with 60 g/l fructose and 2 mg/l 1-naphthalene acetic acid (NAA) and elicitors of 50 mg/l chitosan, 0.1 mM phenylalanine and 100 µM methyl-jasmonate every 5 d. After 45 d of extended culture, intracellular and extracellular paclitaxel levels were analyzed.

Analysis of taxanes (paclitaxel, baccatin III, 10-deacetyl baccatin III) content.

After their separation from the production medium, 0.2 g of cells were weighed, soaked in 4 ml of methanol (Sigma)/dichloromethane (Sigma) (1:1 vol/vol) and sonicated (Branson) for 1 h. The methanol/dichloromethane extract (4 ml) was filtered and concentrated *in vacuo* and subsequently redissolved in 4 ml of dichloromethane and partitioned with 2 ml of water. The latter step was repeated three times and only the dichloromethane fraction was collected. This fraction was concentrated, then redissolved in 1 ml of methanol and centrifuged at 8,000g for 3 min before HPLC analysis. For determining the extracellular paclitaxel concentration, production medium (5 ml) was extracted 3 times with the same volume of dichloromethane. The combined dichloromethane fraction was subsequently concentrated and then redissolved in 0.5 ml methanol. HPLC (nanospace SI-2, Shiseido) with a C18 column (Capcell pak C18 MGII column, 5 µm, 3.0 mm × 250 mm, Shiseido) was used for the analysis. Column temperature was 40 °C and the mobile phase was a mixture of water and acetonitrile (Burdick & Jackson) (1:1 isocratic) at a flow rate of 0.5 ml/min. A UV-VIS detector monitored at 227 nm and the sample injection volume was 10 µl. Authentic paclitaxel, baccatin III, 10-deacetyl baccatin III standard was purchased from Sigma.

Analysis of abietanes (taxamairin A and taxamairin C) content.

After their separation from the production medium, 20 mg of lyophilized cells were weighed, soaked in 4 ml of methanol (Sigma)/dichloromethane (Sigma) (1:1 vol/vol) and sonicated (Branson) for 1 h. The methanol/dichloromethane extract (4 ml) was filtered and concentrated *in vacuo* and subsequently redissolved in 4 ml of dichloromethane and partitioned with 2 ml of water. The latter step was repeated three times and only the dichloromethane fraction was collected. This fraction was concentrated, then redissolved in 1 ml of methanol and centrifuged at 8,000g for 3 min. Then it was filtered through 0.2 µm filter for UPLC analysis. UPLC (Waters) with a C18 column (BEH C18 1.7 µm, 2.1 × 100 mm Waters) was used for the analysis. Column temperature was 40 °C and the mobile phase was a mixture of water and acetonitrile (Burdick & Jackson) flow rate of 0.4 ml/min. Water (solvent A) and acetonitrile (solvent B) as mobile phase with a linear gradient was used: (1 min: 0% B, 13 min: 100% B, 15 min: 100% B, 16.2 min: 0% B, 17 min: 0% B). A UV-VIS detector monitored at 210 nm and the sample injection volume was 2 µl. Authentic taxamairin A and taxamairin C standard were isolated at Unhwa.

Analysis of ginsenosides (ginsenoside F2, gypenoside XVII) content.

Compounds of *Panax ginseng* CMCs were analyzed through HPLC-ELSD (Younglin) and two major peaks were isolated. The two compounds isolated were identified as ginsenoside F2 and gypenoside XVII through LC-MS (Agilent), ¹H NMR, ¹³C NMR and 2D NMR (Varian). For quantification of ginsenoside F2 and gypenoside XVII in *Panax ginseng* CMCs, cultured cells

were separated from the medium and were lyophilized. 100 mg of lyophilized cells were put into 2 ml of methanol (Sigma), vortexed for 5 min, and were extracted for 1 h. Cells were centrifuged at 8,000g for 3 min. After concentration of the supernatant, it was dissolved in 200 μ l of methanol and filtered through 0.2 μ m filter for UPLC analysis. UPLC with a C18 column (BEH C18 1.7 μ m, 2.1 \times 100 mm Waters) was used for the analysis. Column temperature was 40 $^{\circ}$ C and the mobile phase was a mixture of water and acetonitrile (Burdick & Jackson), flow rate of 0.4 ml/min. Water (solvent A) and acetonitrile (solvent B) as mobile phase with a linear gradient was used: (0 min: 0% B, 9 min: 100% B, 11 min: 100% B, 11.2 min: 0% B, 12 min: 0% B). A UV-VIS detector monitored at 203 nm and the sample injection volume was 2 μ l. Standard of gypenoside XVII were isolated by Unhwa. Ginsenoside F2 was purchased from LKT Laboratories.

Microscopy. Light microscopy was undertaken using a model BX41, Olympus. A polarizer for transmitted light, model U-POT, Olympus, was used for TE images. TEs were identified by virtue of their bright fluorescence, due to the presence of lignified secondary cell walls.

Vacuole experimentation was undertaken based on modifications of the methods described previously^{44,45}. Briefly, CMCs, needle- and embryo-derived DDCs were stained with 0.01% (wt/vol) neutral red (SIGMA-ALDRICH) for 3 min. Then, cells were washed with 0.1 M phosphate buffer (pH 7.2) and were observed with an optical microscope (BX41 Olympus) using the same buffer.

LC-MS. Analysis was undertaken using an HP 1100 Series liquid chromatography/HP 1100 Series mass selective detector (Agilent Technologies). Samples (2 μ l) were separated on a PerfectSil Target ODS-3 (4.6 mm \times 150 mm \times 3 μ m) using water (10 mM ammonium acetate): acetonitrile which was isocratic: 50% acetonitrile for 60 min, at 0.4 ml/min flow rate. Mass detection of paclitaxel was by electrospray ionization (ESI) in the positive ion mode. The drying gas was N₂ at 10 l/min, 350 $^{\circ}$ C, 30 p.s.i. The vaporizer was set to

300 $^{\circ}$ C, capillary to 4,000 V. Identification of paclitaxel was accomplished by comparison of retention times and mass with authentic standards.

31. Gamborg, O.L., Miller, R.A. & Ojima, K. Nutrient requirements of suspension cultures of soybean root cells. *Exp. Cell Res.* **50**, 151–158 (1968).
32. Zang, X., Mei, X.-G., Zhang, C.-H., Lu, C.T. & Ke, T. Improved paclitaxel accumulation in cell suspension cultures of *Taxus chinensis* by brassinolide. *Biotechnol. Lett.* **23**, 1047–1049 (2001).
33. Yukimune, Y., Tabata, H., Higashi, Y. & Hara, Y. Methyl jasmonate-induced overproduction of paclitaxel and baccatin III in *Taxus* cell suspension cultures. *Nat. Biotechnol.* **14**, 1129–1132 (1996).
34. Fulcher, N. & Sablowski, R. Hypersensitivity to DNA damage in plant stem cell niches. *Proc. Natl. Acad. Sci. USA* **106**, 20984–20988 (2009).
35. Zhu, Y.Y., Machleder, E.M., Chenchik, A., Li, R. & Siebert, P.D. Reverse transcriptase template switching: A SMART™ approach for full-length cDNA library construction. *BioTech.* **30**, 892–897 (2001).
36. Zhulidov, P.A. *et al.* Simple cDNA normalization using Kamchatka crab duplex-specific nuclease. *Nucleic Acids Res.* **32**, e37 (2004).
37. Margulies, M. *et al.* Genome sequencing in open microfabricated high-density picoliter reactors. *Nature* **437**, 376–380 (2005).
38. Brenner, S. *et al.* Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat. Biotechnol.* **18**, 630–634 (2000).
39. Schmid, R. & Blaxter, M.L. annot8r: GO, EC and KEGG annotation of EST datasets. *BMC Bioinformatics* **9**, 180 (2008).
40. Robinson, M.D., McCarthy, D.J. & Smyth, G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
41. Gentleman, R.C. *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5**, R80 (2004).
42. Bullard, J.H., Purdom, E., Hansen, K.D. & Dudoit, S. Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics* **11**, 94 (2010).
43. Nolan, T., Hands, R.E. & Bustin, S.A. Quantification of mRNA using real-time RT-PCR. *Nat. Protoc.* **1**, 1559–1582 (2006).
44. Kataoka, T. *et al.* Vacuolar sulfate transporters are essential determinants controlling internal distribution of sulfate in *Arabidopsis*. *Plant Cell* **16**, 2693–2704 (2004).
45. Lee, Y. *et al.* The *Arabidopsis* phosphatidylinositol 3-kinase is important for pollen development 1. *Plant Physiol.* **147**, 1886–1897 (2008).

CAREERS AND RECRUITMENT

Third-quarter biotech job picture

Michael Francisco

Biotech and pharmaceutical positions on the three representative job databases tracked by *Nature Biotechnology* (Tables 1 and 2) were mixed in the third quarter of 2010. Compared with the second quarter (*Nat. Biotechnol.* 28, 875, 2010), the two largest biotechs, Monsanto (St. Louis) and Amgen (Thousand Oaks, CA, USA), saw increased job postings whereas most others remained flat or declined. Instrument makers PerkinElmer (Waltham, MA, USA) and Illumina (San Diego) also maintained their hiring pace.

The Boston area was hit hard in particular. Genzyme (Cambridge, MA, USA) announced the reduction of its workforce by 1,000 employees or about 10% over the next 15 months, due to the divestment of its Genzyme

Genetics unit to Laboratory Corp. of America Holdings. And Alnylam Pharmaceuticals (Cambridge, MA, USA) said it would restructure and reduce head count by ~25–30% after completing the fifth and final year of an RNA interference (RNAi) discovery deal with Novartis. Novartis has finalized its selection of 31 disease targets for which it has exclusive rights to discover, develop and commercialize RNAi therapeutics. Alnylam is eligible for development and sales milestones on therapeutics developed against the selected targets. The company said it expects the restructuring to save about \$25 million in operating expenses in 2011.

Table 3 shows additional selected third quarter downsizings within the life sciences industry.

Table 1 Who's hiring? Advertised openings at the 25 largest biotech companies

Company ^a	Number of employees	Number of advertised openings ^b		
		Monster	Biospace	Naturejobs
Monsanto	22,900	4	0	93
Amgen	17,100	52	70	0
Genzyme	12,000	67	1	0
CSL	10,300	0	0	0
Life Technologies	9,000	21	35	0
PerkinElmer	8,800	61	0	0
Bio-Rad Laboratories	6,600	9	14	0
bioMerieux	6,300	18	0	0
Millipore	6,100	19	14	0
Novozymes	5,122	0	0	0
IDEXX Laboratories	4,800	3	0	0
Biogen Idec	4,750	12	100	1
Shire	3,875	78	0	0
Gilead Sciences	3,852	0	29	0
WuXi PharmaTech	3,773	0	0	0
Biocon	3,673	0	0	0
Qiagen	3,500	0	0	0
Cephalon	3,026	0	0	0
Celgene	2,813	15	17	0
Actelion	2,200	5	3	0
Biotest	1,834	0	3	0
Illumina	1,781	34	59	16
Amylin Pharmaceuticals	1,500	4	2	0
Vertex Pharmaceuticals	1,427	43	79	1
Elan Pharmaceuticals	1,321	8	7	0

^aAs defined in *Nature Biotechnology's* survey of public companies (28, 793–799, 2010).

^bAs searched on Monster.com, Biospace.com and Naturejobs.com, 13 October 2010. Jobs may overlap.

Table 2 Advertised job openings at the ten largest pharma companies

Company ^a	Number of employees	Number of advertised openings ^b		
		Monster	Biospace	Naturejobs
Johnson & Johnson	119,200	>1,000	5	5
Bayer	106,200	72	26	2
GlaxoSmithKline	103,483	0	0	1
Sanofi-aventis	99,495	0	0	2
Novartis	98,200	104	91	10
Pfizer	86,600	7	103	64
Roche	78,604	25	39	4
Abbott Laboratories	68,697	12	26	1
AstraZeneca	67,400	146	0	0
Merck & Co.	59,800	0	9	0

^aData obtained from *MedAdNews*. ^bAs searched on Monster.com, Biospace.com and Naturejobs.com, 13 October 2010. Jobs may overlap.

Table 3 Selected biotech and pharma downsizings

Company	Number of employees cut	Details
Adolor	30	Restructured and reduced head count by 30% to about 80 to save cash.
Amylin Pharmaceuticals	60	Cut about 4% of its workforce in a "few discrete functions," including R&D, but cuts do "not signal a move away from research."
Cypress Bioscience	123	Will restructure and reduce head count by 86% to 20 after it discontinued co-promotion of fibromyalgia drug Savella (milnacipran) in the US with partner Forest Laboratories.
Febit Group	~48	Restructured and reduced head count by about 60% to 30–35 to focus on its blood-based microRNA biomarker discovery services.
Gilead Sciences	120	Will close its HCV and HBV research facility in Durham, NC by the end of the year and move the operations to Foster City, CA, USA.
Merck & Co.		Announced plans to eliminate eight R&D and eight manufacturing facilities over the next 2 years as part of an ongoing restructuring following last year's acquisition of Schering-Plough, along with a 15% head count reduction.
NicOx	22	Closed its US headquarters after FDA issued a complete response letter for an NDA for naproxinod to treat signs and symptoms of osteoarthritis, resulting in head count reductions of 18% to 102.
Phenomix	45	Has cut most of its employees and is winding down operations after failing to find a new partner in North America to develop and commercialize diabetes candidate dutogliptin.
Santarus	120	Will reduce head count by about 37% to about 200, the majority from sales, as the company is ceasing promotion of its prescription Zegerid (omeprazole/sodium bicarbonate) due to impending generic competition.
Stem Cell Therapeutics	NA	Expects head count reductions to cut costs after NTx-265 to treat stroke missed its primary endpoint in a phase 2b trial.
Vectura Group	85	Will restructure R&D operations and reduce head count by 31% to 190 to focus on partnered programs and branded generics. The cuts include the proposed closure of a Nottingham, UK R&D facility.

HCV, hepatitis C virus; HBV, hepatitis B virus; NDA, new drug application; NA, not available. Source: BioCentury.

Michael Francisco is a Senior Editor, *Nature Biotechnology*



PEOPLE



Privately held NKT Therapeutics (Waltham, MA, USA) has announced the appointment of **Barbara Finck** (left) to the position of chief medical officer. She was most recently senior vice president of R&D and chief medical officer at Osprey Pharmaceuticals. Prior to that, she served as the vice president of clinical development at Eos Biotechnology and then at Protein Design Labs following its acquisition of Eos.

"We are delighted to add someone with Barbara's background and depth of experience," says Robert Mashal, CEO of NKT

Therapeutics. "This important addition to the NKT Therapeutics senior management team further advances our goal of building a pipeline of first-in-class, NKT, cell-directed therapeutics to treat autoimmune and inflammatory diseases, cancer, infectious diseases and dermatitis."

Agile Therapeutics (Princeton, NJ, USA) has named **Al Altomari**, a former senior manager at Johnson & Johnson who has been serving as Agile's chairman, as president and CEO, succeeding **Thomas Rossi**. Rossi, who joined Agile as president and CEO in 2004, will serve as a scientific advisor to the company. Altomari was most recently CEO of Barrier Therapeutics.



Paul R. Billings (left) has been named chief medical officer of Life Technologies (Carlsbad, CA, USA), a newly created position. Billings brings extensive expertise and clinical experi-

ence, most recently serving as director and CSO of the Genomic Medicine Institute at El Camino Hospital. He currently serves as a member of the US Department of Health and Human Services Secretary's Advisory Committee on Genetics, Health and Society, where he helps shape policy in the field of genomic medicine, and has been a founder or CEO of companies involved in genetic and diagnostic medicine, including GeneSage, Omicia and CELLective Dx.

Nick Glover has been named CEO and a director of YM BioSciences (Mississauga, Ontario, Canada). **David Allan** has relinquished his post as CEO but will continue to serve as chairman of the company. Glover joined YM in June 2010 as president and COO.

SpectraScience (San Diego) has announced that **Jim Hitchin** has resigned his position as

chairman of the board but will remain as CEO of the company. The board has elected **Mark McWilliams** to assume the duties of chairman. McWilliams, CEO of Medipacs, has served as a SpectraScience director since 2004.

Pacgen Biopharmaceuticals (Vancouver) has announced the election of **John Hsuan** to its board of directors. Hsuan has incubated and co-founded more than 20 public companies and currently serves as chairman of NCTU Venture Capital, Maxima Capital Management and Faraday Technology. He also currently serves as the emeritus vice chairman of United Microelectronic, a leading global semiconductor company.

Pharming Group (Leiden, The Netherlands) has announced the appointment of **Karl Keegan** as CFO. Keegan formerly served in the same capacity at Minster Pharmaceuticals and has more than 15 years of industry experience, including as a biotech analyst at several investment banks.

Vaccinogen (Frederick, MD, USA) has appointed **Michael L. Kranda** as president and CEO. He brings more than 20 years of experience developing biotech platforms, products and companies and serves on the board of directors of several companies, including PTC Therapeutics. Kranda was most recently CEO, president and director of Anesiva. Prior to joining Anesiva, Kranda was managing director of Vulcan Capital.

Genetic analysis software maker Geospiza (Seattle) has announced the appointment of

Bruce Montgomery to its board of directors. Montgomery has over two decades of experience in pharma and biotech management, most recently serving as senior vice president of respiratory therapeutics at Gilead Sciences. Previously, he had founded and served as CEO of Corus Pharmaceuticals until its sale to Gilead.

Advaxis (North Brunswick, NJ, USA) has named **Robert Petit** to the newly created position of vice president of clinical operations and medical affairs. He joins Advaxis from Bristol-Myers Squibb, where he served as the US medical strategy lead for the ipilimumab program, director of medical strategy for new oncology products and director of global clinical research.

Alexander Polinsky has been named to the board of directors of AtheroNova (Irvine, CA, USA). He is managing partner of Maxwell Biotech Venture Fund. He also co-founded Alanex and served as its CSO until it was acquired by Agouron Pharmaceuticals in 1997. After Agouron's acquisition by Pfizer, Polinsky served in several capacities, including vice president, head of discovery technologies and CEO of The Pfizer Incubator.

Privately held NGM Biopharmaceuticals (S. San Francisco, CA, USA) has named **William J. Rieflin** as CEO and a member of the board of directors. Rieflin brings over 20 years of industry experience to NGM, most recently as president and director at XenoPort. Prior to that, he was executive vice president, administration, CFO, general counsel and secretary for Tularik, which was acquired by Amgen in 2004.

Thomas Werner has been named president and CEO of Accera (Broomfield, CO, USA). He was previously managing director and senior vice president of GlaxoSmithKline in Germany and had served as CEO of the GSK Foundation. In March 2009, Werner was elected to the Accera's board of directors.

Altheos (S. San Francisco, CA, USA) has appointed **Barbara Wirostko** chief medical officer. Wirostko was previously senior medical director and the glaucoma medical team leader at Pfizer where she was responsible for pipeline strategy in glaucoma.