THE NEW ATLANTIS
A JOURNAL OF TECHNOLOGY & SOCIETY

Who Owns the Genome?

Misha Angrist and Robert M. Cook-Deegan

Like other ambitious public endeavors, the Human Genome Project was prone to the occasional spectacle. On June 26, 2000, President Bill Clinton (with British Prime Minister Tony Blair hovering via satellite) stood at the White House flanked by two men: Francis Collins, director of the National Human Genome Research Institute and leader of the public consortium that sequenced the human genome, and J. Craig Venter, CEO and founder of the biotech firm Celera Genomics. Celera's mission was to sequence the human genome better and faster than its government-funded rival. It aimed to sell access to genomic information as well as the tools to interpret it, with an eye to "big pharma" and other biotechnology companies looking for a treasure trove of new drug targets.

Both the public and private sequencing efforts garnered their share of headlines, but Celera's genome-era business model failed miserably. In hindsight this is not surprising: Why would drug companies pay millions of dollars for sequence data that would be publicly available just months later? Within 18 months of the White House event, Venter was deposed as CEO of Celera, and the company refashioned itself as a drug discovery and development firm. The genome wars officially ended in April 2005, when Celera announced it would donate its sequence data to the public databases.

The race to map the human genome seemed to embody two rival ways of doing science: the public pursuit of human knowledge paid for with public dollars and the private pursuit of useful knowledge paid for with private investments. Each way of doing science has its purist adherents, but in reality the relationship between "public" science and "private" science is a complicated thicket, especially when it comes to the brave new world of genomics. Nation by nation, laboratory by laboratory, regulation by regulation, we are still trying to figure out who "owns" the genome, what the owners actually own, and how best to balance the pursuit of knowledge, the allocation of rewards, and the development of life-improving biotechnologies. Innovation continues apace, but no one really knows how close we are to the optimum policies. The system works, but

Misha Angrist *is science editor at the Duke Institute for Genome Sciences and Policy.* ***Robert M. Cook-Deegan*** *is director of the institute's Center for Genome Ethics, Law, and Policy.*

no one knows how well, because no one can address the crucial question: compared to what?

Scientific Ideals and Market Incentives

Thirty years ago, the sociologist Robert Merton articulated the norms that distinguish the scientific enterprise—much of it straightforward, but still worth restating. Science is a communal, social process. Research works best when those engaged in scientific endeavors are dispassionate about the outcome and when scientific honors are bestowed based on merit alone. Scientific claims are expected to be original—to build on but extend beyond the work of others. And scientific research should be governed by an underlying skepticism, meaning that all claims are tentative, theory-dependent, subject to empirical verification and rigorous scrutiny by other scientists.

What links these norms is the notion of "openness," an assumption that there should be a "scientific commons" where information and materials are made freely available to fellow scientists at low or no cost. Traditionally, academic research, including in genomics, has come closest to this Mertonian ideal. Different disciplines, institutions, and individuals vary widely in practice, but the ideal itself is (supposedly) shared by all true scientists.

Private research works differently, even when the experiments themselves are similar. Fame and fortune are driving forces in both the commercial and scientific-academic worlds, but the balance between them is different: fame in science depends on the esteem of a small community of experts; fortune depends on making knowledge useful to the broader public in the form of technology. Financial incentives, secrecy, management of information, and publication are part of both worlds, but publication and data-sharing are bedrock to science, while management of information and returning value for investment are the dominant goals in private R&D, and quite properly so.

While biotechnology and pharmaceutical companies may not share their data as quickly or openly as academic researchers, however, the patent system often exposes private-sector research to the same scrutiny as the peer-review process. Patents may be ridiculed as nefarious tools of capitalism, but patents come with a duty to *disclose*, albeit on an *ex post facto* basis. Patents are, in a crucial sense, the *public* governance of private enterprise. That is why they are called patents—meaning "open"—in direct contrast with trade secrets. Moreover, private R&D is populated

largely by those with academic training, and some companies, including some very successful ones such as Genentech, actively promote open publication once patent rights have been secured. In this way, the norms of the science commons are partially transposed from academic to commercial laboratories, even as the reward structures in private versus public genomics remain quite distinct.

Since the 1970s, government policy has encouraged university scientists to reap the commercial benefits of their research. The landmark Bayh-Dole Act of 1980 spurred academic institutions to pursue patents on federally-funded research and to license their inventions to private firms for commercialization. Earlier that same year, the U.S. Supreme Court paved the way for patents on genetically modified organisms, ruling in favor of a scientist who had developed a bacterium by selective breeding (*Diamond v. Chakrabarty*).

Many years later, we are still sorting out the legacy of these policies and court decisions. Public genomics and private genomics are more intermingled than ever. Universities face decisions about whether to purchase commercial licenses in order to conduct academic research. Companies support academic researchers in order to preempt rivals by making certain areas of knowledge public and thus unpatentable. Some argue that an excess of exclusive patents restricts genomic research, while others argue that the backlog of patent applications is slowing down genomic research in a maze of fruitless regulation. Within this maze, however, a few patterns seem to have emerged—some showing how the private and the public realms can serve one another, others suggesting that the system stands in need of serious reform.

My Virus Is Your Virus

The SARS virus offers a recent, telling example of how intellectual property (IP) can help and hinder research and development simultaneously. The virus first appeared in China's Guangdong province in late 2002. By spring 2003, the SARS genome was already sequenced, and the sequence data was quickly made public. Without question, data sharing greatly expedited the characterization of the virus and the ability to diagnose infection. Sharing information saved lives by saving time. Within six months, vaccines were being tested, and a DNA chip harboring the complete SARS genome was available for both research and screening.

At the same time, the institutions that sequenced the SARS genome filed patent applications, and these patents could prove important if a private

sector partner someday needs to supply capital for developing a vaccine. In the case of SARS, several groups (Hong Kong University, a Dutch medical center, a Canadian public health agency, and the U.S. Centers for Disease Control) initially filed patent applications; the latter two of these groups contended they were patenting for defensive purposes—that is, to prevent others from locking up SARS data and rights. Yet in a recent *Bulletin of the World Health Organization*, the Dutch patent-applicants expressed concern that the SARS patents may deter the development of downstream products such as vaccines. Remarkably, intellectual property can evoke ambivalence even among those who hope to have it.

Unlike the genome wars, we don't yet know how the SARS story will end. The verdict on the value of SARS patents depends on imponderables that will not be apparent for years. The various parties could agree to share their patents with one another, if not the world, just as they have already shared the virus's DNA sequence. Or the SARS rights could remain proprietary. Most likely, few people will pay much attention so long as SARS stays off the global radar screen, in which case sorting out the patent situation will probably take longer and cost more than sequencing the virus in the first place. The fear of mass death will not be there to focus the mind, so to speak.

But it is also possible that SARS could return, or that some new pathogen—like bird flu—will raise similar questions about how IP can be deployed in a way that protects public health while keeping financial incentives in place. And ironically, those situations when cooperation is most needed—such as during an epidemic—are also the times when the most money may be at stake. Disease-related patents are most valuable, after all, if and when the disease in question becomes a true public health concern.

Private Interest, Public Good

The SARS case demonstrated that information sharing can be a force for social good, and that patents might permit such sharing while still preserving incentives for subsequent development (though such sharing is hardly guaranteed). A different pattern is what might be called "strategic altruism"—that is, when companies support the expansion of public knowledge for the sake of private interests.

In the realm of genomics, there are at least two notable examples of such strategic altruism. In 1994, pharmaceutical giant Merck bankrolled an effort to detect regions of the genome that harbored active, protein-coding genes. The company was worried about its freedom to operate in

future years, since upstart companies like Human Genome Sciences and Incyte Genomics were filing patent applications on hundreds of thousands of short snippets of genes. Such patents threatened to lock up exclusive rights to make, use, or sell the human genes of which the snippets were only a part, creating thousands of "toll booths" along the way to producing final products. In response, Merck funded Washington University in St. Louis to produce the same partial gene sequence information quickly and publicly. Merck helped create new public information in order to preemptively defeat rival intellectual property.

A second example of strategic altruism came in 1999, when a consortium of drug companies sank $30 million into an effort to identify and make public millions of genetic variants that help identify particular diseases and drug responses. Like Merck's strategic response to the partial sequencing initiative, the consortium was spurred by the actions of smaller genomics firms (such as Celera) who were seeking to develop proprietary databases of these bits of DNA.

These two cases illustrate the topsy-turvy world of genomics in action, with big companies working with universities and nonprofits to bolster the public domain. Was it pure self-interest, driven by market demands and exigencies? Probably. But the net consequences were largely positive both for the advance of human knowledge and for the development of new products.

Interestingly, the incentive for "big pharma" to pay for such public research only existed because of the patent system itself—that is, because smaller biotechnology companies were working to sequence potentially patentable bits of DNA. Governments and philanthropists may support science for its own sake or as a benefit to mankind, but private R&D is driven largely by the belief that it will produce valuable IP, particularly patented inventions based on DNA and its uses. No doubt many corporate scientists are public-spirited and interested in knowledge for its own sake. But companies are only likely to make significant investments in public knowledge when there is a private incentive to do so. This is not greed, but reality.

The Many Routes to Knowledge

Knowledge, after all, is never free. It requires time, talent, and resources—and thus investments that offer the possibility of economic returns. Even as the market value of the fifteen largest genomics firms dropped five-fold from 2000 to 2002 (from $50 billion to $10 billion), their R&D spending continued to rise, despite a difficult economic environment.

These companies expect that some of this research will lead to useful knowledge, and the general public expects that the incentive system driving such research will produce goods and services of commercial and social value. We recognize that progress always requires risk and usually entails failures along the way, and that risk-takers need the possibility of gain if they are to endure those research avenues (indeed, the vast majority) that go nowhere fast or nowhere at all.

Yet the market model is not the only model for advancing science, and the system of private-sector incentives exists alongside a more truly "public" genomics. In 1996, as the Human Genome Project was shifting into a higher gear, a meeting of the major genome sequencing players led to an agreement to disclose DNA sequence data on a daily basis. While it would be a few years before the data were transferred daily to the GenBank database where they could be most useful, data were nonetheless available much earlier than they would have been otherwise. Moreover, the agreement succeeded in linking a network of "early users" to the large-scale sequencing centers, thereby speeding the entire human genome sequencing enterprise. Signatories were also asked not to patent DNA sequences unless they could provide an inventive step beyond the sequence data itself. Rapid disclosure increased the efficiency of the overall effort by avoiding duplication.

These practices were modeled in part on successes in the nematode research community. The "worm network" was a powerful model of rapid scientific advances flowing from open science norms. It was "a hub-and-spoke" way of doing science, with a few big central laboratories doing cutting-edge biology while still rapidly sharing information and tools with a large number of smaller laboratories all over the world. Of course, it would be naïve to equate the *modi operandi* of labs working on worms with those working on humans—the technological and economic stakes are different, because the relevance of human genetics to curing human disease is far more significant, to say the least. But the worm experience was a powerful demonstration of public science at its best, leading to a series of substantive contributions to basic biology and even a few Nobel prizes.

In general, the greatest value produced by academic genomics research is derived from freely available information—information that can be used by scientists, clinicians, and companies for subsequent innovation. The social value of that information utterly dwarfs the financial value to universities of licensing fees attached to the inventions arising from the research. The recent debate on freely available "open access publication" and exemplars such as the Public Library of Science and other new approaches to scientific publishing suggest that, at least in the aca-

demic sector, publication remains the primary way of keeping score and creating value through data sharing. Preserving this culture of productive openness—alongside, not instead of, the culture of market science—is one of the big policy challenges in the years ahead.

File Now, Fight Later

So how might we begin to think about reconciling these two approaches to advancing science—patents with openness, innovation with access? Consider yet another example of genome technology, in which both the possibilities and limits of the current IP system are being played out in real time: RNA interference (RNAi).

RNAi is the means by which double-stranded RNA binds to its DNA counterpart and silences the gene encoded by that stretch of DNA. RNAi is a natural process first observed in plants 15 years ago, but its molecular basis was not understood until 1998. The timing turned out to be fortuitous: just as genome scientists were overwhelmed with thousands upon thousands of genes of unknown function—that is, the fruits of the Human Genome Project—along came a technique that would give immediate insight into what those genes actually did. By using RNAi to silence genes of interest, researchers can often discover and understand gene function.

Elucidating gene function via RNAi is already an entrenched tool in molecular biology. But we don't yet know whether this approach will ultimately break through as a therapeutic tool, allowing us to knock out mutant genes in patients with a minimum of toxicity. Best case, RNAi will eventually account for billions of dollars and a significant segment of the drug market. Worst case, it will prove to be more hype than hope—mired in false promise *à la* gene therapy.

To judge from the tangled mass of RNAi IP, many people are bullish on the therapeutic possibilities. As of late 2005, roughly 2,000 RNAi-related patents were awaiting judgment from the U.S. Patent and Trademark Office (PTO). Yet the number of U.S. RNAi patents actually issued by the PTO can still be counted on one hand. This absurd imbalance points to some systemic flaws in the way the PTO functions. A new technology, coupled with a limited number of patent examiners and a gold-rush mentality, has led to the current feeding frenzy. For many researchers and research institutions, standard procedure is to file first and ask questions later. Consequently, RNAi patent applications will take years to sort out, and litigation by some of the two dozen companies and various universities with a stake in RNAi will likely take years.

Despite the gridlock, one hears relatively little bellyaching from RNAi practitioners either in industry or academe—in part because RNAi is a pre-clinical and relatively immature technology, but also because of the emerging legal and regulatory environment. Two recent and somewhat contradictory court decisions are central. In *Madey v. Duke* (2003), a federal appeals court held that research scientists are not exempt from patent infringement—so that when a patent-holder comes prowling through the labs of academia looking to get paid, no longer can a researcher get away with saying, "Wait! I was only doing research!" Then, in *Merck v. Integra* (2005), the U.S. Supreme Court decided that companies *could* get away with infringing others' patents without paying, so long as their research was related to new drugs intended for FDA approval. These mixed signals from the courts have created a confusing and conflicted environment for research.

Initially, researchers feared that the *Madey* decision would have a chilling effect, as private genomics firms pursued patent-infringing academic scientists. But because of the *Merck* decision and the fact that there has been no real fallout from *Madey*, academic genome researchers feel more protected from litigation. And researchers using RNAi methods to develop drugs for FDA approval might be shielded by the *Merck* decision (although the downside, of course, is that developers of RNAi tools stand to lose major licensing revenue). At any rate, many patent holders are opting not to litigate, if only for practical reasons: Small biotechs rarely have the resources to chase down infringers and seek redress via the courts. To paraphrase one RNAi company CEO, for now it's much easier to license to one's competitors than it is to block them. Moreover, suing research laboratories for patent infringement is not worthwhile when the damages are limited to fees for reagents used in research—tiny sums by comparison to the hoped-for gigabucks from future RNAi treatments. We'll likely only see legal action when the stakes get high enough, and the future of this technology will only take shape once litigation in informative cases is complete and the contradictions of *Madey* and *Merck* are ironed out.

A Menu of Solutions

Not surprisingly, many stakeholders are eager to improve the current IP system—corporations worried about the costs of litigation and the backlog of patents, academic scientists worried about preserving the culture of free intellectual exchange, policymakers looking to maximize the economic return on public research dollars.

One approach to improving the way we use and share knowledge is gaining steam. As documented by Arti Rai, James Boyle, and Jerry Reichman at Duke Law School, a number of public funding agencies, individual scientists, and even some private drug and biotechnology companies have begun to embrace a more collaborative ethos. Much of this behavior is occurring in the realm of bioinformatics software, an area where licensing tends to yield little in the way of revenue. Software is often distributed free of charge and on a so-called "copyleft" basis, whereby all users are free to modify and redistribute it; subsequent recipients inherit the same freedom to modify and redistribute, but cannot impose restrictions on future users. Other software has been made freely available so long as its use is acknowledged, and some software has been made available with no restrictions at all, meaning it *can* be incorporated into commercial products.

A number of large-scale genome database projects have also embraced a "copyleft-style" ethic. The publicly funded International Haplotype Mapping Project ("HapMap") published its first catalogue of human genetic variation in October 2005; the goal of the project is to link patterns of genetic variation with complex diseases. The HapMap released genotype data immediately (just as the public human genome sequencing consortium pledged to do in 1996). In exchange for access, HapMap users agreed not to file product patent applications if they relied, even in part, on HapMap data—a condition being relaxed now that the catalogue is public and the data are largely complete.

Other organic approaches are emerging from the scientific trenches. Science Commons (ScienceCommons.org) seeks ways for scientists to exchange information in a way that minimizes "unnecessary transaction costs." One of its goals is to craft private, voluntary material transfer agreements written in plain English, without the bureaucratic bells and whistles that have come to characterize agreements for exchange of materials, particularly between industry and academic users.

Of course, it would be naïve to expect genomic IP problems to resolve themselves entirely from the bottom up. Some policy solutions will need to be imposed, and improving the patent application and award process is a good place to start. As legal scholars Lori Andrews and Jordan Paradise have pointed out, the patent system vis-à-vis genes seems broken. The PTO is understaffed and its backlog is enormous. Moreover, as the system is presently constructed, there is a financial incentive for examiners to award patents: give a patent, get a bonus. Yet a substantial number of the thousands of gene patents already awarded fail to meet the PTO's own

criteria of novelty and non-obviousness. One possible improvement would be to remove the incentive to grant patents by finding an alternative metric of compensation for examiners. Another possibility is to allow patents to be contested for a short period after they are issued, by "opposition" proceedings, mimicking current policy in Europe. This might root out some patents that should not have been granted in the first place.

Licensing agreements can also avert excessive patent litigation. Patent pools like those in the electronics industry, whereby patent owners broadly license their patents to one another, would be another formal mechanism to minimize costs. In the event that voluntary patent pools or licensing do not take hold, the government could mandate compulsory licensing in certain limited cases where the public interest is at stake—as it did to break patent logjams in aircraft engines, radio broadcast, and naval propulsion technologies.

Taken together, these bottom-up initiatives and top-down policies offer simply a patchwork approach to improving the way genomics is done. And given a realm with so much complexity, so many interests, and so many unknowns, tinkering with the parts is probably the best we can do. But it is also possible that the gathering enthusiasm for "open and collaborative" research, even in the private sector, signals an inflection point. Perhaps we have moved beyond the impassioned rhetoric of public *versus* private; perhaps we no longer regard the human genome as either "the common heritage of all mankind" immune from IP rights or as a Wild West for speculative patents and endless court fights.

The ultimate fate of genomic research—who (if anyone) owns, pays for, and innovates with genomic information—won't be known for decades. But we are increasingly moving beyond the two-dimensional modes of thinking that characterized the early days of biotechnology. Legions of genome scientists (Craig Venter among them) now promote patenting and commercialization in some areas, such as protein-based drugs, while simultaneously promoting open science and expressing hostility to restrictive patents in others, such as software and raw data. If we are going to maximize the benefits of the genomic revolution, we need an intellectual property system capable of making such fine distinctions—always remembering that public policy and modern science move forward in a messy, incremental, imperfect way.