

SYNTHETIC BIOLOGY: CAUGHT BETWEEN PROPERTY RIGHTS, THE PUBLIC DOMAIN, AND THE COMMONS

ARTI RAI* & JAMES BOYLE**

Novel artificial genetic systems with twelve bases instead of four.¹ Bacteria that can be programmed to take photographs² or form visible patterns.³ Cells that can count the number of times they divide.⁴ A live polio virus “created from scratch using mail-order segments of DNA and a viral genome map that is freely available on the Internet.”⁵ These are some of the remarkable, and occasionally disturbing, fruits of “synthetic biology,” the attempt to construct life starting at the genetic level. In terms of their scale and ambition, these efforts go beyond traditional recombinant DNA technology. Rather than simply transferring a pre-existing gene from one species to another, synthetic biologists aim to make biology a true engineering discipline. In the same way that electrical engineers rely on standard capacitors and resistors, or computer programmers rely on modular blocks of code, synthetic biologists wish to create an array of modular biological parts that can be readily synthesized and mixed together in different combinations. MIT has a “Registry of Standard Biological Parts [which] supports this goal by recording and indexing biological parts that are currently being built and offering synthesis and assembly services to construct new parts, devices, and systems.”⁶ Systems, devices, parts, and DNA represent descending levels of complexity – systems consist of devices, and devices consist of parts composed of DNA. The idea behind a registry of parts is that these parts can, and should, be recombined in different ways to produce many different types of devices and systems. Although the Registry currently contains physical DNA, its developers believe that, as DNA synthesis technology becomes increasingly inexpensive,⁷ the Registry will be composed largely of information and specifications which can be executed in synthesizers just as semiconductor chip designs are executed by fabrication firms.

Synthetic biology has already produced important results, including more accurate AIDS tests and the possibility of unlimited supplies of previously scarce drugs for malaria.⁸ Proponents hope to use synthetic organisms to produce not only medically relevant chemicals but also a large variety of industrial materials, including biofuels such as hydrogen and ethanol.⁹ At the same time, synthetic biology has engendered numerous policy concerns. From its inception, commentators have raised issues ranging from bioethical and environmental worries to fears of bioterrorism – indeed, the Central Intelligence Agency released a report in 2003 called “The Darker Bio-Weapons Future” that explicitly referred to the dangers posed by the possibility of genetically engineered “super-viruses”¹⁰

There is, however, one area that has been largely unexplored until this point – the relationship of synthetic biology to intellectual property law. Two key issues deserve further attention. First, synthetic biology, which operates at the confluence of biotechnology and computation, presents a particularly revealing example of a difficulty that the law has frequently faced over the last 30 years – the assimilation of a new technology into the conceptual limits around existing intellectual property rights, with possible damage to both in the process. There is reason to fear that tendencies in the way that the law has handled software on the one hand and

biotechnology on the other could come together in a “perfect storm” that will impede the potential of the technology. Second, synthetic biology raises with remarkable clarity an issue that has seemed of only theoretical interest until now. It points out a tension between different methods of creating “openness”. On the one hand, we have intellectual property law’s insistence that certain types of material remain in the public domain, outside the world of property. On the other, we have the attempt by individuals to *use* intellectual property rights to create a “commons,” just as developers of free and open source software use the leverage of software copyrights to impose requirements of openness on future programmers, requirements greater than those attaching to a public domain work. Intellectual property policy, at least in the United States, specifies things that cannot be covered by intellectual property rights, such as abstract ideas or compilations of unoriginal facts, precisely to leave them “open” to all – the public roads of the intellect.¹¹ Yet many of the techniques of open source *require* property rights so that future users and third parties will be bound by the terms of the license. Should we rethink the boundary lines between intellectual property and the public domain as a result?

THE PERFECT STORM: FLAWED BIOTECH LAW MEETS FLAWED SOFTWARE LAW?

Intellectual property law has already had difficulty incorporating the revolutionary technologies from which synthetic biology draws inspiration – biotechnology and computers. In the area of biotechnology, the U.S. Court of Appeals for the Federal Circuit (which hears most patent appeals) has tended not to enforce the patent law requirement that inventions be “nonobvious” to the ordinary scientist working in the area. Years after methods for cloning genes became routine and widely known, the Federal Circuit continues to treat the gene products of such methods as patentable.¹² On the Federal Circuit’s reasoning, what matters is not whether a practicing biologist would find a particular invention obvious but, rather, *per se* rules about nonobviousness developed for chemical inventions in the mid-20th century.¹³ So one major part of the technological terrain into which synthetic biology must fit – biotechnology – has already proven difficult for intellectual property law to manage.

While biotechnology has mainly posed difficulties for patent law, computers have posed both copyright and patent problems. Copyright covers original works of expression. It explicitly excludes works that are functional. Patent law covers inventions that are useful, novel and non-obvious – functionality is a requirement, not an impediment. However, it had traditionally been understood to exclude formulas and algorithms. Thus software – a machine made of words, a set of algorithmic instructions devoted to a particular function – seemed to fit neither the copyright nor the patent box. It was too functional for copyright, too close to a collection of algorithms and ideas for patent. What’s more, certain economic aspects of software, including its high propensity to display “network effects” (increased utility from product use based on increase in number of others using the product) led scholars believe that both copyright and patent were ill-suited to encourage innovation without discouraging competition. Several *sui generis*, or custom-made, intellectual property regimes were proposed as an alternative.

In the event, as a result of statements by Congress and actions by the courts, software ended up being covered by *both* copyright and patent – a result that most scholars thought was far from ideal. Moreover, the refusal by at least some members of the Federal Circuit to allow

patent examiners to use unwritten information to determine whether a particular patent application is obvious¹⁴ may have a disproportionate impact on computer-related inventions. Because much knowledge in the field of computer technology is not written down in journal articles, it may be hard for a patent examiner to find specific written references testifying to information that is generally known. Additionally, many scholars have argued that the Federal Circuit allows unduly broad patents to issue in the area of software.¹⁵

How does this history of intellectual property law's struggles to deal with software and biotechnology bear on synthetic biology? Consider the following patent, issued by the PTO on August 10, 2004. The patent, number 6,774,222, issued to the U.S. Department of Health and Human Services ("HHS"), is entitled "Molecular Computing Elements, Gates, and Flip-flops." This patent covers using the combination of nucleic-acid binding proteins and nucleic acids to set up data storage as well as logic gates that perform basic Boolean algebra. As the patent document notes, the invention could be used not only for computation but also for complex ("digital") control of gene expression. The broadest claim, claim 1, does not limit itself to any particular set of nucleic-acid binding proteins or nucleic acids. Many types of molecular computing and control of gene expression are likely to be covered by such a patent. Moreover, the claim uses language that would cover not only the "parts" that performed the Boolean algebra but also any device and system that contained these parts. Such a patent would seem effectively to patent algebra, or the basic functions of computing, when implemented by the most likely genetic means. It is difficult even to imagine the consequences of an equivalent patent on the software industry. What is the likelihood that this foundational patent, or one similar to it, would hold up in court?

Given the low nonobviousness threshold that the Federal Circuit has set in the area of biotechnology, there is some possibility that the court would apply a similarly low threshold here. But even if the HHS invention were viewed as more similar to software or electronics than to biotechnology, the nonobviousness threshold could still be low. This is because, as noted earlier, the Federal Circuit is reluctant to allow unwritten knowledge to be used in determining nonobviousness. So even if, at the time the HHS invention was made, individuals working in the field knew that many computing functions could readily be performed using DNA, this unwritten knowledge might not be factored into the nonobviousness determination. Additionally, to the extent that the HHS patent were viewed as software, it would probably not be considered too broad. Notably, the HHS patent is not unique in its breadth. A Stanford University patent claims the use of computer system to simulate the operation of a biochemical network, at least for a specified period of time.¹⁶

Considerable historical evidence, including evidence from virtually every important industry of the 20th century, suggests that broad patents on foundational research can slow growth in the industry.¹⁷ In this regard, it may be particularly instructive to contrast the proprietary situation in the nascent area of synthetic biology with that of computer hardware, computer software, and biotechnology in their infancy. In the area of computer hardware, the specter of broad patents loomed large until government action forced licensing of the AT&T transistor patent as well as patents obtained by Texas Instruments and Fairchild Instruments on integrated circuits. As for software, it was already a robust industry before software patents became available. Biotechnology's foundational technologies – monoclonal antibodies and

recombinant techniques – either were not patented or were made available widely at reasonable cost. Synthetic biology may be coming of age under different circumstances, at the juncture of two technologies with which the law was already struggling.

In addition to the problem of broad foundational patents, there is the possibility of a plethora of narrower patents, some of which may fall within the scope of the foundational patents. For example, scientists at Boston University have filed somewhat narrower patents that claim the use of DNA to produce specific gene regulation mechanisms such as a multi-state oscillator; a genetic toggle switch; and an adjustable threshold switch.¹⁸ MIT and Sangamo have patents on various types of DNA-binding proteins. At least in the area of information technology,¹⁹ there is considerable evidence that patent thickets²⁰ or “anti-commons”²¹ create difficulties for subsequent researchers above and beyond those created by foundational patents. This is because many products in information technology represent combinations of dozens if not hundreds of patented parts. To the extent that work in synthetic biology is similarly likely to rely on combinations of many parts, the thicket or anti-commons difficulty may be quite salient. Not only does a crowded patent landscape create the possibility of “hold-up” by a previously unknown patent holder who emerges only after others have invested large sums of money in the area of the patented invention, but to the extent that patent rights holders rely upon reach-through royalties to secure revenue, standard economic theory predicts that product output by the improver will be suboptimal.²² Moreover, while firms that work in information technology have sometimes succeeded in pooling patents, particularly patents around industry standards, efforts at patent pooling do not always succeed in addressing problems of inefficient royalty-stacking.²³ Such efforts have also been stymied by failure on the part of participating firms to disclose relevant patents.²⁴ In any event, because synthetic biology encompasses not only information technology but also biotechnology, the absence of any successful patent pools in the life sciences is cause for concern.

A SYNTHETIC BIOLOGY COMMONS?

These intellectual property concerns have not gone unnoticed. The MIT scientists involved with the Registry of Standard Biological Parts are sufficiently concerned that they have created a “Bio-Bricks Foundation” that might serve to coordinate a synthetic biology “commons.” The idea of a synthetic biology commons draws inspiration, in part, from the prominence of the open source software model as an alternative to proprietary software. Like software, synthetic biology aims to be information-based and modular. Indeed, the synthetic biologist might argue that what she does is comparable to software programming – the only difference is that synthetic biologists programs with four bases (As, Ts, Cs, and Gs) while ordinary software programmers use 0s and 1s. So the analogy to open source software is hardly far-fetched.

Unlike proprietary software developers, open source software producers make their source code freely available for improvement, modification, and redistribution. Certain types of open source licenses also have a “commons-expanding” aspect: these “copyleft” licenses not only make source code freely available, but they also require those who distribute improvements to the source code to make the improvements available on the same terms.²⁵ Copylefted

software relies heavily on the existence of property rights – specifically, copyright in the source code. Because of this copyright, users of the copylefted software necessarily use it subject to the terms of the license.

Synthetic biologist might argue that strings of DNA bases are comparable to source code and that DNA strings could therefore also be covered by copyright. The difficulty with this line of reasoning is that software itself fits poorly into copyright's categories. Congress indicated a desire that software be covered by copyright, but left it to the courts to work out the method of doing so. As developed by the courts, copyright protection in software is thin – for example, source code is generally protected against verbatim copying. Even with source code, moreover, if the code is entirely dictated by functional concerns or has become an industry standard, it may not be protected by copyright *at all*.

Where does this leave synthetic biology? There are two major obstacles. First, unlike software, the products of synthetic biology are not discussed as copyrightable subject matter in the statute. Thus a court that wished to find that material copyrightable would have to do so by analogy. Second, even if courts were willing to make such an analogy, there are the internal restrictions of copyright law, which does not cover functional articles or methods of operation, and requires expressive choices. Thus even for a court willing to take the first step and analogize synthetic biology to software, there would be further obstacles to overcome. As a matter of legal doctrine, the answer might depend upon the type of synthetic biology involved. For example, the construction of DNA sequences using base pairs that do not exist in nature might allow significant room for expressive choice. Such DNA sequences might be protected by copyright, at least against verbatim copying. However, most synthetic biologists working today, including those at MIT, are working within the confines of the existing genetic code. This code constrains the expressive choices that they make, making copyright protection less likely. Beyond formal legal doctrine lies a set of policy concerns. With patent rights clearly available, courts and Congress might be reluctant to layer on an entirely new kind of property right, for fear that such rights would hurt rather than help innovation. The fact the question of copyrightability arises in the attempt to create a research commons should not change the conclusion. While the goal is a laudable one, the boundaries of the public domain should not be altered to enable a particular initiative.

Thus, in the case of synthetic biology, the ability to invoke copyright is by no means clear. An obvious alternative is patents. One example of a patent-based commons is that created by the group Biological Innovation for an Open Society (BIOS). BIOS is using patent protection on a few key plant gene transfer technologies to force licensees to put improvements to those technologies into the commons.²⁶ Although some have suggested that the BIOS approach could raise concerns about antitrust and patent misuse,²⁷ the concern should be relatively small given BIOS's mission to expand the commons and the relatively permissive, rule of reason-based approach taken by contemporary antitrust law. The more pressing problem for purposes of projects like the MIT Registry – which contains more than two thousand standardized parts – is that a patent-based approach may be quite expensive. A single patent can cost tens of thousands of dollars to secure.

Of course, to the extent that a few broad patents might effectively cover many of the parts in the Registry, the patent option becomes more plausible. For example, a patent comparable in

breadth to the HHS patent noted above might cover many Registry parts. In this scenario, the Registry would essentially be exploiting flaws in the current patent system for commons-expanding purposes. The difficulty in this scenario would be to identify an area of inventive territory that was quite broad but nonetheless not suggested either by prior broad patents or by information already in the public domain.

Alternatively, the Registry might try to attract statements of non-assertion by other patentees, on the model of recent statements by IBM, Sun Microsystems, and other firms that they will not assert their patents against anyone working on open source software. Indeed, the fact that many synthetic biology patents are currently held by academic and government institutions may make such statements of assertion a real possibility. Non-assertion statements would certainly be useful in providing those who are working on the MIT Registry comfort in moving forward. More generally, to the extent that institutions with synthetic biology patents vowed not to assert their patents against academic researchers, such a move would be a salutary development. Non-assertion statements are not, however, a property right. In order to secure a property right, the owners of the MIT Registry would need a license with explicit permission to sublicense. Moreover, patents licensed to the Registry would have to cover, at least in some fashion, parts that were important for maintaining and expanding the commons.

Another alternative for securing an expanding commons might rely on some kind of contract, such as a “clickwrap” license over the Bio-Bricks. This contractual alternative does not require an underlying property right. Instead, the contract simply imposes conditions as part of the price of access. One problem with such contracts is that they bind only those who receive the technology from the entity imposing the terms. Attempts to prevent leakage to those not bound by the terms of the contract can require strict restrictions on information dissemination. For example, for some time the publicly funded International HapMap project (a database of human genetic variation) used a click-wrap license. This license required those who sought access to single nucleotide polymorphism (“SNP”) data to refrain from combining it with their own proprietary SNP data in order to seek product patents on haplotypes (collections of SNPs). In order to prevent leakage of the data outside the confines of this clickwrap license, to those who would then have no obligation to the HapMap commons, the license required those who sought the data to refrain from disseminating to anyone who had not signed on to the license. Conventional publication of the data was not possible. This condition is no longer imposed because it is believed that the database has reached a sufficient density to be self-sustaining and to defeat subsequent patent claims. But the old requirements indicate one of the difficulties of the ‘clickwrap’ approach; the comparative weakness of the contractual restraints paradoxically requires extremely broad restrictions on dissemination.

Finally, legislative proposals might create *sui generis* property rights mechanisms for protecting BioBricks data. Indeed, the European Union currently has *sui generis* protection of data. The evidence suggests, however, that strong property rights protection is likely to hinder rather than promote innovation.²⁸ A recent draft of the proposed Access to Knowledge treaty offers an alternative *sui generis* approach: under this approach, member countries would adopt legislation protecting “qualifying open databases” from patents on certain types of improvements for a specified period of time.²⁹ Various commentators affiliated with the Access to Knowledge proposal have also suggested the possibility of “social patents” legislation: under this approach, a

type of patent right could be secured at low or no cost, but it could not be used for exclusionary commercial purposes. Although these *sui generis* alternatives are quite intriguing, and certainly an improvement over ordinary property rights in databases, securing new legislation is a difficult, uncertain, and slow route.

We close with one overarching observation. Copyleft licenses, which lead to the formation of an ever-expanding commons, have worked well – even brilliantly – in the software context. Copyright licenses have produced well-functioning code, and they have also constrained the threat posed by copyright and patent, particularly when such intellectual property could be attached to an incipient industry standard. Will they work as well in synthetic biology? There is reason for some caution. Intellectual property rights are relatively unimportant *as incentives* at any stage in the production of copyleft software. They are important mainly for the leverage they give to the licensor. But synthetic biology might be different. Though the uses of synthetic biology are by no means limited to biomedicine, at the end of some biological chains of innovation will lie the expensive development and commercialization of a drug. While taking a drug all the way through FDA-mandated clinical trials may not cost as much as drug companies claim, it does cost hundreds of millions of dollars. Whether patent rights are the best incentive mechanism for purposes of eliciting pharmaceutical R&D is not a question we can address here. Suffice it to say that our current system of financing pharmaceutical innovation relies heavily on these rights. There is no direct equivalent in the world of free software. If a copyleft condition – however drafted and imposed – did attach to some of synthetic biology parts, care would have to be taken in the design of the system, lest the result was to make it impossible for that technology to be developed into a patented therapy. The BIOS licenses, which restrict the copyleft condition to improvements on the enabling technology, and do not constrain patenting on transgenic plant products, provide an interesting model. But the distinction between enabling technology and product may be easier to make in a situation like that faced by BIOS, where the enabling technology in question has a relatively clear innovation trajectory, both in terms of improvement to the technology itself and in terms of production of end products.

In the mean time, the decision, already implemented, of the MIT Registry to place its parts into the public domain certainly provides important protection against the threat of patents clogging innovation in the synthetic biology space. Placing parts into the public domain not only makes the parts unpatentable, but it undermines the possibility of patents on trivial improvements. In the end, a public domain strategy comparable to that employed by the public Human Genome Project may not be ideal, but it is certainly a good start.

Table 1

Strategy	Advantages	Disadvantages*
Patents (20 years from time of patent application to exclude others from use of invention)	Clear property right basis for copyleft license (license that requires improvements to be distributed freely)	Expensive (approx. \$25,000 per patent in U.S. for complex inventions)
Copyright (attaches immediately upon creation; exclusive right to copy and improve that lasts for 70 years after author's death, or if work of corporate authorship, 95 years from publication)	Clear property right basis for copyleft license; inexpensive	Legal basis for assertion of copyright unclear – no explicit basis in the Copyright Act and some theoretical arguments against.
Contract (terms vary)	Inexpensive	Copyleft license requires strict limits on information dissemination
<i>Sui generis</i> (one of a kind) legislation (“Open” databases; “social patents”)	Narrowly tailored to problem	Legislative solutions are difficult and slow.

*In all of the copyleft approaches, there are line-drawing issues. One has to be very careful regarding precisely what material is covered by the requirement.

* Professor of Law, Duke Law School; Faculty Associate, Institute for Genome Sciences & Policy

** William Neal Reynolds Professor of Law, Duke Law School. The authors gratefully acknowledge the support of the National Human Genome Research Institute and the Department of Energy (P50 HG003391-02). We thank Drew Endy, Stephen Maurer and four anonymous reviewers for their comments.

¹ Benner, SA (2003) Act natural. *Nature* 421: 118-118.

² Levskaia, A, Chevalier, AA, Tabor, JJ, Simpson, ZB, Lavery, LA, et al. Synthetic biology: Engineering *Escherichia coli* to see light. *Nature* 438: 441-442.

³ Basu, S, Gerchman, Y, Collins, CH, Arnold, FH, and Weiss, R (2005) A synthetic multicellular system for programmed pattern formation. *Nature* 434: 1130-1134.

⁴ Endy, D (2005) Foundations for engineering biology. *Nature* 438: 449-453.

⁵ Ball, P (2004) Synthetic biology: Starting from scratch. *Nature* 431: 624-626.

⁶ <http://parts.mit.edu/> (Registry of Standard Biological Parts) Accessed 2006 Oct 3

⁷ Carlson, R (2003) The Pace and Proliferation of Biological Technologies. *Biosecurity and Bioterrorism* 1:1-12.

⁸ Martin, VJ, Pitera, DJ, Withers, ST, Newman, JD, and Keasling, JD (2003) Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nature Biotechnology* 21:796-802.

⁹ <http://www.syntheticgenomics.com/corporate.htm> (website for Synthetic Genomics, Inc.) Accessed 2006 Oct 3

¹⁰ Central Intelligence Agency, *The Darker Bioweapons Future* Directorate of Intelligence 3rd November 2003, OTI SF 2003-108.

-
- ¹¹ Within patent law, the scope of the “abstract ideas” exception to patentability has progressively been narrowed by the Court of Appeals for the Federal Circuit. Although the Supreme Court was poised to decide this question in a recent case, a majority of the Court subsequently decided that the issue had not squarely been presented. *LabCorp. v. Metabolite Corp.*, 126 S.Ct. 2921 (2006) (per curiam)
- ¹² *In re Deuel*, 51 F.3d 1552 (Fed. Cir. 1995).
- ¹³ Rai, AK (1999) *Intellectual Property Rights in Biotechnology: Addressing New Technology*. *Wake Forest Law Review* 34, 827-; *see also* *In re Fisher*, 421 F.3d 1365, 1382 (Rader, J., dissenting) (noting this point).
- ¹⁴ *In re Sang-Su Lee*, 277 F.3d 1338 (Fed. Cir. 2002).
- ¹⁵ Burk, D and Lemley, M (2002) *Is Patent Law Technology-Specific?* *Berkeley Technology Law Journal* 17:1155-
- ¹⁶ U.S. Patent No. 5,914,891
- ¹⁷ Merges, RP and Nelson, RR (1990) *On the Complex Economics of Patent Scope* *Columbia Law Review* 90, 839- (1990)
- ¹⁸ U.S. Patent Nos. 6,737,269; 6,841,376; 6,828,140.
- ¹⁹ The situation in biotechnology is less clear. Compare Walsh, JP, Cho, C, and Cohen WM *View from the bench: Patents and materials transfers*. *Science* 309: 2002-2003 with Murray, F & Stern, S *NBER Working Paper* 11465.
- ²⁰ C. Shapiro, 1 *Innovation Policy and the Economy* 119 (A. B. Jaffe et al., eds. 2001).
- ²¹ Heller, M and Eisenberg, RS (1998) *Do Patents Deter Innovation? The Tragedy of the Anticommons in Biomedical Research*. *Science* 280: 698-
- ²² Lemley, M & Shapiro, C *Patent Holdups and Royalty Stacking* (working paper 2006)
- ²³ *Id.*
- ²⁴ *In the Matter of Rambus, Inc.*, Docket No. 9302, Opinion of the Commission, August 2, 2006
- ²⁵ See <http://www.gnu.org/licenses/licenses.html#LicenseList> (discussing GNU General Public License and other “copyleft” licenses)
- ²⁶ www.bios.org
- ²⁷ Burk, D & Boettiger, S (2004) *JIBL* 1, 221 (2004); Feldman, R (2004) *Minnesota Journal of Law, Science, and Technology* 5: 117 (2004)
- ²⁸ Commission of the European Communities, DG Internal Market and Services Working Paper, 12 December 2005
- ²⁹ Treaty on Access to Knowledge, Article 5-6, draft of 9 May 2005