ValpoScholar **Valparaiso University Law Review**

Volume 46
Number 1

*pp.*1-41

Fall 2011

DNA Copyright

Andrew W. Torrance

Articles

DNA COPYRIGHT

Andrew W. Torrance*

TABLE OF CONTENTS

I. INTRODUCTION[1]

Each year since 1980, on a Friday in January, the Massachusetts Institute of Technology ("MIT") has hosted the MIT Mystery Hunt.[2] To win this competition, a team must solve a number of challenging puzzles whose solutions reveal where on the MIT campus to find a special coin. The first team to locate the coin wins the Mystery Hunt and is awarded the privilege of designing the next Mystery Hunt.

In 2005, the clues of one of the puzzles, entitled "Shotgun Wedding," involved 11 nucleotide sequence fragments, each approximately 1,000 nucleotides in length.[3] To solve this puzzle, it was necessary to understand that the fragments resulted from "shotgun sequencing," a

---

* Andrew W. Torrance is a Professor in the School of Law, a Research Associate at the Biodiversity Institute, and a Docking Faculty Scholar at the University of Kansas. He received a Ph.D. in biology from Harvard University and a J.D. from Harvard Law School. The author would like to thank his research assistants, Dr. James McParland and Dr. Sparkle Ellison, for their invaluable assistance with research for this article.

1 The author first presented his ideas about DNA copyright in November 2008 in a speech entitled "Synthesizing Law for Synthetic Biology," at *Biolaw 2.0: Law at the Frontiers of Biology*, at the University of Kansas School of Law. Subsequently, the author presented his proposals regarding DNA copyright in March 2009 in "Synthetic Biology—Law and the Next Open Source Hardware," at the MIT Innovation Lab, MIT Sloan School of Management; in May 2009 in "Synthesizing Law for Synthetic Biology," at the OECD Collaborative Mechanisms for Intellectual Property Management Workshop, OECD Headquarters, Paris, France; in June 2009 in "Synthesizing Law for Synthetic Biology," at the Biotech Summer Institute at Drake University and Pioneer Hi-Bred, Des Moines, Iowa; in November 2009 in "The Legal Possibilities of Synthetic Biology," at the MIT Innovation Lab, MIT Sloan School of Management; in September 2010 in "Open Biological Innovation—From Patents to Commons to Copyright to Open Source," at the Berkeley Open Innovation Forum, Berkeley Law Schools, Berkeley, California; in January 2011 in "Synthetic Biology Meets the Law," at the American Association of Law Schools Annual Meeting, San Francisco, California; and in March 2011 in "Open Biological Innovation: From Patents to Commons to Copyright to Open Source," at the Valparaiso University Law Review Symposium on Bioethics, Law, and Synthetic Biology, Valparaiso University School of Law, Valparaiso, Indiana. The author has previously published articles about gene patents Andrew W. Torrance, *Gene Concepts, Gene Talk, and Gene Patents*, 11 MINN. J.L. SCI. & TECH. 157, 157–91 (2010) [hereinafter, Torrance, *Gene Concepts, Gene Talk, and Gene Patents*], and Andrew W. Torrance, *Patent Rights and Civil Wrongs: The ACLU Lawsuit*, 8 BIO-IT WORLD 11 (2009); synthetic biology, Andrew W. Torrance, *Synthesizing Law for Synthetic Biology*, 11 MINN. J.L. SCI. & TECH. 157, 642–48 (2010) [hereinafter, Torrance, *Synthesizing Law for Synthetic Biology*]; and genomics, Andrew W. Torrance, *Family Law and the Genomic Revolution*, 79 U. MO. KAN. CITY L. REV. 271 (2011) that draw on some common research sources.

2 *MIT Mystery Hunt*, MIT.EDU, http://web.mit.edu/puzzle/www/index.html (last visited June 27, 2011). Thank you to James Grimmelmann for bringing this to my attention on the IPProfs listserv (October 14, 2010).

3 Jed Goldstone, *Shotgun Wedding*, MIT.EDU, http://web.mit.edu/puzzle/www/05/setec/shotgun_wedding (last visited June 27, 2011).

method of determining the sequential nucleotides in a stretch of deoxyribonucleic acid ("DNA"). The fragments had to be aligned "contiguously," and the ends of "contig[s]" translated into their corresponding amino acids. Done properly, this resulted in the following instruction: "SEEKGENEFRAMES." Finally, the longest "open reading frame[ ]," when translated into its corresponding amino sequence, yielded the message "MYANSWERISSEPARATECHECKSANDTHERESTISFILLER."[4] Armed with this clue, teams could progress to the next stage in the Mystery Hunt.

An intriguing feature of the Shotgun Wedding puzzle is that the DNA sequences it employed were designed by Jed Goldstone not to be precursors for the synthesis of a polypeptide that might be useful in treating disease, conferring a useful trait on a crop plant, or carrying out an industrial process. Rather, Goldstone simply used the genetic code of nucleotides adenine ("A"), guanine ("G"), thymine ("T"), and cytosine ("C"), and the amino acids (for example, methionine ("M"), tyrosine ("Y"), and alanine ("Ala")) encoded by triplet codons of these nucleotides to create a short intelligible message. Goldstone was an author, his medium was DNA, and his products were original and expressive literary works. Although Goldstone employed A, G, T, and C as symbols to represent nucleotides, he might instead have chosen to compose his work using actual nucleotides decipherable through routine chemical sequencing techniques. Either way, the original work of authorship that Goldstone created in a tangible medium of expression using sequences of nucleotides and amino acids is eligible for copyright protection. In fact, even DNA sequences that code for functional polypeptides or RNAs may qualify for copyright protection to the extent that function does not dictate structure, and expression is not unduly constrained.

The idea of DNA copyright is not new.[5] As long ago as 1982, Irving Kayton concluded that copyright protection is available for DNA.[6] Since then, several other authors have similarly concluded that DNA constitutes subject matter eligible for copyright.[7] However, as Rebecca

---

4 *Id.*

5 In this article, the discussion explicitly focuses on the applicability of copyright law to DNA. However, this discussion is also germane to RNA, a similar and related nucleic acid, and polypeptides, which are functionally related to both DNA and RNA.

6 Irving Kayton, *Copyright in Living Genetically Engineered Works*, 50 GEO. WASH. L. REV. 191, 191–92 (1982).

7 Dan L. Burk, *Copyrightability of Recombinant DNA Sequences*, 29 JURIMETRICS J. 469, 531–32 (1988–89); Jorge A. Goldstein, *Copyrightability of Genetic Works*, 2 BIO/TECHNOLOGY 138

Eisenberg observed in 1990, "copyright protection for DNA sequences has failed to make its mark outside the scholarly literature."[8] Since then, this situation may have begun to change.

In 2002, an organization called the DNA Copyright Institute began to advertise a service involving the "copyrighting" of a person's genome to guard against infringement or misappropriation by others.[9] However unlikely it is that an existing naturally occurring genome could constitute a work of authorship, the idea of DNA copyright has begun to gain traction, especially with respect to partially or fully synthesized genes or other DNA sequences.[10] Improvements in methods of gene sequencing and gene synthesis have transformed the prospect of designer DNA from laborious and unpredictable to routine and certain. Furthermore, the burgeoning field of synthetic biology is founded, at least in part, on the promise of deliberately engineering genes, cells, and organisms *de novo*. In fact, the use of DNA copyright has already begun in industry. Illumina, Inc., a biotechnology firm whose genome sequencing machines lead the genomics industry,[11] produces DNA molecules to be used with its machines, and views their molecules as works of genetic authorship. Moreover, Illumina explicitly asserts copyright protection over some of its DNA sequences. To illustrate, the following is a letter sent to an Illumina customer:

> Dear Customer,
> This communication is in response to your request for particular oligonucleotide sequences for use with the Illumina Genome Analyzer and associated assays. Below please find the oligonucleotide sequences that we can make available to you. This communication is solely for your use and should not be distributed outside your institution.

(1984); Donna Smith, Comment, *Copyright Protection for the Intellectual Property Rights to Recombinant Deoxyribonucleic Acid: A Proposal*, 19 ST. MARY'S L.J. 1083, 1096–1108 (1988).

8 Rebecca Eisenberg, *Patenting the Human Genome*, 39 EMORY L.J. 721, 721 n.3 (1990).

9 Peter Huck, *How the Rich and Famous will Fight to Stay Unique*, SUNDAY TRIB. (Mar. 28, 2010), http://tribune.maithu.com/article/2001/oct/14/how-the-rich-and-famous-will-fight-to-stay-unique/.

10 In this article, "gene" and "DNA sequence" are often used interchangeably, where appropriate. However, DNA sequences constitute the broader, more inclusive category. Genes fall within one particular category of DNA sequences, while there are many other types of DNA sequences that are not genes.

11 *See Fact Sheet*, ILLUMINA.COM, http://www.illumina.com/Documents/company/IlluminaCorporateSheet_050410.pdf (last visited July 17, 2011) (claiming Illumina, Inc. is a "leading developer, manufacturer, and marketer of life science tools and integrated systems for large-scale analysis of genetic variation and function.").

> The oligonucleotide sequences are protected by copyright which is owned by Illumina. Illumina allows you to reproduce the oligonucleotide sequences for use with the Illumina Genome Analyzer and associated assays. Additionally, Illumina realizes customers may need to make alterations to the oligonucleotide sequences that are necessary to allow use with the Illumina Genome Analyzer and associated assays. Thus, Illumina allows you to make such necessary alterations to the oligonucleotide sequences but only for use with the Illumina Genome Analyzer and associated assays. Illumina grants you no other rights to use, reproduce or otherwise disclose the oligonucleotide sequences. Alteration or modification of the oligonucleotide sequences for use with non-Illumina products is not allowed.
>
> If you reproduce the oligonucleotide sequences for viewing within your institution, the following copyright notice must remain affixed to the sequences:
>
> > Oligonucleotide sequences © 2006 Illumina, Inc. All rights reserved. Illumina customers may reproduce and create derivative works of the oligonucleotide sequences but only for use with the Illumina Genome Analyzer and associated assays. All other uses are strictly prohibited.
>
> If you reproduce these oligonucleotide sequences for viewing outside your institution (e.g. journal publication), you must affix the following copyright notice to the sequences:
>
> Oligonucleotide sequences © 2006 Illumina, Inc. All rights reserved.[12]

Like Goldstone, Illumina designed the oligonucleotides referenced in the above letter. The nucleotide sequences of the primers and adapters are not copies of genomic DNA sequences; they are synthetic sequences

---

[12] Letter from Illumina for Customers (Oct. 13, 2010), *available at* http://www.bioinfo.uh.edu/IMDSC/Release_of_Oligo_Sequences_Letter_for_Customers1.pdf. Thank you to Jessica Sibley for bringing this to my attention on the IPProfs listserv.

not found in nature.[13] In fact, it is probably important that they not be genomic DNA sequences, because, if they were, they might hybridize with complementary DNA sequences loaded onto Illumina sequencing machines. Rather, these synthetic sequences must employ designs that ensure their compatibility with machines sold by Illumina. The Illumina customer letter indicates that the company considers its authored oligonucleotides to be protected by copyright. Although these Illumina sequences may possess more functionality than those composed by Goldstone, Illumina is correct in its assumption that authored synthetic DNA sequences are eligible for copyright protection.[14]

This Article suggests that DNA—especially synthetic DNA—constitutes eligible subject matter for copyright protection under the Copyright Act. Although DNA has long been copyrightable, in theory, the movement towards synthetic DNA in biotechnology has further strengthened existing arguments in favor of DNA copyright. Section II of this Article illustrates some of the features of DNA that suit it for copyright protection by tracing the conceptual evolution of DNA from factor to program. Section III suggests the usefulness of DNA copyright through a discussion of the recent rocky road down which gene patents have been traveling. Section IV sketches the rise of synthetic biology as a distinct field. Section V outlines why DNA is eligible for copyright protection, considers implications of DNA copyright, and then discusses benefits that might accrue to society under a DNA copyright regime, including those flowing from fair use provisions and the fostering of open source biology. The Article concludes by suggesting that DNA copyright: (1) already exists; (2) provides an alternative to DNA patenting; and (3) may provide a number of societal benefits in terms of biological innovation and improved societal access to the fruits of such innovation.

II. EVOLVING CONCEPTS OF DNA

Few concepts in science have undergone such rapid and complete transformation as the concept of the unit of heredity. At various times in

---

[13] Rochelle Dreyfuss, IPProfs Listserv (Oct. 14, 2010).

[14] The United States Copyright Office appears to have acknowledged the copyrightability of DNA in 1987. *See* J. SIGALOS, INTELLECTUAL PROPERTY PROTECTION FOR BIOTECHNOLOGY INNOVATIONS 13 (1987) (stating the "Copyright Office advises drawing of DNA nucleotide sequence bearing copyright notice sufficient for copyright"). *But see* OFFICE OF TECH. ASSESSMENT, OTA-BA-370, NEW DEVELOPMENTS IN BIOTECHNOLOGY: PATENTING LIFE—SPECIAL REPORT 43 (1989) (noting that the Copyright Office's unofficial position that nucleic acid sequences are not copyrightable, citing Peter R. Bahn & Steven J. Hultquist, *Engineered Proteins as Intellectual Property*, GENETIC ENGINEERING NEWS, Mar. 1987, at 18–19).

history, what is often called the "gene" was believed to be a spirit, a liquid, or a particle. Now, the gene is generally conceived of as a stretch of DNA capable of encoding information, which, in turn, is capable of acting as a template for constructing complementary sequences of DNA, RNA (directly), or polypeptides (indirectly). Apart from some viral RNA genomes, genes in nature are usually composed of DNA.

Often only a fraction of the DNA in genes encodes "exons" and is expressed as mRNA and by extension, polypeptide products.[15] In fact, non-expressed DNA, or "introns," in mammals, insects, and birds comprise about 80% of the nucleotide length of a gene.[16] Moreover, the vast majority of the human genome—about 98%—does not encode polypeptides.[17] Noncoding and apparently nonfunctional DNA has been dismissively termed "junk DNA." Much of this noncoding DNA may eventually be discovered to possess functionality, but it appears likely that much of the DNA in the genomes of humans and other organisms may lack direct function.

As the understanding of DNA has evolved over the last century, so have the economic and legal treatments of DNA. From the beginnings of the biotechnology industry in the 1970s, many have sought to attach intellectual property protection to DNA and related molecules. The protection offered by trade secrecy depends on the extent to which the molecule to be protected is disclosed by the product or service of which it forms a part, but is agnostic about how exactly science conceives of that molecule. The aptness of patent protection for DNA increased as the conception of genes underwent a transition from inchoate factors to discrete molecules. Patent protection for DNA seemed assured in 1980 by the landmark Supreme Court decision in *Diamond v. Chakrabarty*.[18] However, patent protection for DNA isolated from genomic sources has lately been put into doubt by the Southern District of New York's decision in *Ass'n for Molecular Pathology v. U.S. Patent & Trademark Office*, and remains in doubt following the Second Circuit's ruling on the case.[19] As the conception of DNA shifted again from a mere physical molecule to a repository of information and instructions, copyright became a more

---

15 BENJAMIN LEWIN, GENES IX 45 (2008).

16 *Id.*

17 Greg Elgar & Tanya Vavouri, *Tuning in to the Signals: Noncoding Sequence Conservation in Vertebrate Genomes*, 24 TRENDS GENETICS 344, 345 (2008).

18 Diamond v. Chakrabarty, 447 U.S. 303, 309 (1980).

19 Ass'n for Molecular Pathology v. U.S. Patent & Trademark Office, 702 F. Supp. 2d 181 (S.D.N.Y. 2010) (holding that patents on DNA that indicate breast cancer susceptibility were invalid). On July 29, 2011, the Federal Circuit reversed the Southern District of New York's decision in part and affirmed in part. Ass'n for Molecular Pathology v. U.S. Patent & Trademark Office, 2011 WL 3211513 (Fed. Cir. July 29, 2011).

promising source of intellectual property protection. To understand the transition in intellectual property eligibility, it is important to trace how the conception of genes and DNA has evolved.

A. Factor

One of the signal achievements of Charles Darwin was the idea that evolution occurs by means of "descent with modification."[20] However, Darwin did not determine the specific mechanism by which modifications were passed along from ancestors to descendants. Rather, as suggested by the title of his masterwork, *The Origin of Species by Means of Natural Selection*, he identified natural selection as a higher order cause of evolutionary change. What Darwin did contribute was speculation about the existence of "gemmules" as the specific units of heredity.[21]

At about the same time, an Austrian monk named Gregor Mendel had formed a hypothesis that plant traits were passed from parents to offspring by means of "*Elemente*."[22] Through carefully controlled experiments on pea plants, Mendel not only demonstrated the efficacy of his *Elemente* hypothesis, but also elucidated the basic rules of heredity.[23] Mendel's foundational research came to prominence only after it was published in English in 1901.[24]

Neither Darwin's "gemmules" nor Mendel's "*Elemente*" survived as a description of the unit of heredity. That distinction went instead to the "gene," a word coined by Wilhelm Johannsen and inspired by "pangens," a word coined by Hugo de Vries.[25] The word "gene" rapidly rose to dominance as a description of the fundamental unit of heredity.

B. Particle

The unit of heredity gradually underwent a transition from a vague factor of unknown form to a discrete and physical object. In the 19th century, August Weismann proposed the existence of "determinants," which he imagined to be particles of "a definite chemical, and above all,

---

[20] CHARLES DARWIN, ON THE ORIGIN OF SPECIES BY MEANS OF NATURAL SELECTION 456 (Harvard Univ. Press 1966) (1859).

[21] *See* EVELYN FOX KELLER, THE CENTURY OF THE GENE 2 (2000) (explaining Darwin's impact on the understanding of genes).

[22] *Id.* at 19.

[23] *See generally* Gregor Mendel, *Experiments in Plant Hybridisation*, 26 J. ROYAL HORTICULTURAL SOC'Y 1 (1901) (proving that certain pairs of differentiating characters, the germ-cells of a hybrid, or cross-bred, are *pure*, being carriers and transmitters of either the one character or the other, not both).

[24] *Id.*

[25] KELLER, *supra* note 21, at 1–2.

molecular composition."[26] His views were echoed by his contemporary, Hugo de Vries, who insisted that "[j]ust as physics and chemistry go back to molecules and atoms, the biological sciences have to penetrate to these units in order to explain, by means of their combinations, the phenomena of the living world."[27] In addition to clarifying the mechanisms of heredity, understanding the particulate nature of hereditary units facilitated the growth of experimental genetics.

In the "Fly Room" at Columbia University, Thomas Hunt Morgan used the fruit fly (*Drosophila melanogaster*) as a model organism for experimental genetics.[28] Among other accomplishments, Morgan was able to demonstrate that genes reside on chromosomes, and in his book entitled *The Theory of the Gene*, he described the orientation of genes on chromosomes as "beads on a string."[29] Due largely to Morgan, whose research won him the 1933 Nobel Prize in Medicine, "genes became generally viewed as discrete, stable, independently segregating units of inheritance lined up along a chromosome, an image captured by the most commonly invoked metaphor—that of 'beads on a string.'"[30] However, if they can be viewed as particles, genes are very peculiar particles that encode remarkable amounts of information.

C. Molecule

By the middle of the 20th century, the idea that genes consisted of discrete particles had become firmly entrenched in biology. However, the precise structural and functional nature of these particles had not yet been elucidated. Both polypeptides and nucleic acids, such as DNA, had been proposed as the carriers and determinants of hereditary traits. The scales tipped decisively in favor of DNA following the Avery-MacLeod-McCarty experiment. In 1943, two Canadians, Oswald Avery and Colin MacLeod, and an American, Maclyn McCarty, experimentally determined that bacterial genes were composed of DNA, which "must be regarded not merely as structurally important but as functionally active in determining the biochemical activities and specific

---

[26] August Weismann, *The Continuity of the Germ-Plasm as the Foundation of a Theory of Heredity*, *in* ESSAYS UPON HEREDITY AND KINDRED BIOLOGICAL PROBLEMS 167, 168 (Edward B. Poulton et al. eds., 1891).

[27] HUGO DE VRIES, INTRACELLULAR PANGENESIS 13 (C. Stuart Gager trans., The Open Court Publishing Co. 1910) (1889).

[28] *See* JAMES D. WATSON ET AL., MOLECULAR BIOLOGY OF THE GENE 12–14 (Beth Wilbur et al. eds., 6th ed. 2008).

[29] THOMAS HUNT MORGAN, THE THEORY OF THE GENE 24 (1928).

[30] Leonie Moyle, *Most Ingenious: Troubles and Triumphs of the Century of Genes*, 17 BIOLOGY & PHIL. 715, 715–16 (2002).

characteristics of pneumococcal cells."[31] Alfred Hershey and Martha Chase later confirmed this discovery in blender experiments using bacteria and bacteriophage.[32] The vague concept of a gene as a physical particle had been married to a specific molecule: DNA.

D. Sequence

In 1953, James Watson and Francis Crick announced that they had discovered the physical conformation of DNA: two individual DNA molecules wind around each other in antiparallel to form a double helix.[33] As Watson and Crick explained:

> In the double helix, the two DNA chains are held together by hydrogen bonds . . . between pairs of bases on the opposing strands . . . . This base pairing is very specific: the purine adenine only base-pairs to the pyrimidine thymine, whereas the purine guanine only base-pairs to the pyrimidine cytosine. In double-helical DNA, the number of A residues must be equal to the number of T residues, whereas the number of G and C residues must likewise be equal . . . . As a result, the sequence of the bases of the two chains of a given double helix have a complementary relationship, and the sequence of any DNA strand exactly defines that of its partner strand.[34]

This announcement "convinced biologists not only that genes are real molecules but also that they are constituted of nothing more mysterious than deoxyribonucleic acid."[35] Building on the insight of the double helix, Seymour Benzer was soon able to demonstrate that genes were linear stretches of DNA sequence.[36]

Once DNA had been understood to comprise linear arrays of nucleotides, much research was devoted to determining the exact nucleotide sequences of DNA molecules of interest. At first, these efforts were laborious and inefficient. Early sequencing efforts relied on restriction enzymes, which are proteins capable of snipping a DNA

---

31 Oswald T. Avery et al., *Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types*, 79 J. EXPERIMENTAL MED. 137, 155 (1944).

32 *See generally* A. D. Hershey & Martha Chase, *Independent Functions of Viral Proteins and Nucleic Acid in Growth of Bacteriophage*, 36 J. GEN. PHYSIOLOGY 39 (1952).

33 WATSON ET AL., *supra* note 28, at 22.

34 *Id*.

35 KELLER, *supra* note 21, at 3.

36 *Id*. at 52.

molecule at sites characterized by specific patterns of nucleotides. Restriction enzymes can be employed to recognize locations in the genome hosting particular nucleotide motifs.[37] However, by the 1970's, sequencing methods had improved so substantially that one could identify the precise pattern of individual nucleotides in a DNA molecule. These methods worked as follows:

> The underlying principle of DNA sequencing is based on the separation, by size, of nested sets of DNA molecules. Each of the DNA molecules starts at a common 5' end, and terminates at one of several alternative 3' endpoints. Members of any given set have a particular type of base at their 3' ends. Thus, for one set, the molecules all end with a G, for another a C, for a third an A, and for the final set a T. Molecules within a given set (e.g., the G set) vary in length depending on where the particular G at their 3' end lies in the sequence. Each fragment from this set therefore indicates where there is a G in the DNA molecule from which they were generated.[38]

Allan Maxam and Walter Gilbert invented one of the leading methods, known as "Maxam-Gilbert sequencing," which used "four different regimens of chemical treatment that cause [radiolabeled DNA molecules] to break preferentially at Gs, Cs, Ts, or As."[39] Frederick Sanger invented the other leading method, called the "[Sanger] chain-termination method," which used the enzyme DNA polymerase to create complementary copies of fragments of the DNA molecule being sequenced.[40] Gilbert and Sanger shared the 1980 Nobel Prize in chemistry "for their contributions concerning the determination of base sequences in nucleic acids."[41] Further improvements in sequencing have made the determination of DNA sequences from all biological sources routine. Even entire genomes, such as those of humans, have been

---

37 ANTHONY J. F. GRIFFITHS ET AL., INTRODUCTION TO GENETIC ANALYSIS 343–60 (2005).

38 WATSON ET AL., *supra* note 28, at 753.

39 *Id.* at 754; *see* Allan M. Maxam & Walter Gilbert, *A New Method for Sequencing DNA*, 74 PROC. NAT'L ACAD. SCI. U.S.A. 560 (1977).

40 WATSON ET AL., *supra* note 28, at 754.

41 *See* Bo G. Malmstrom, *The Nobel Prize for Chemistry (1980)*, *in* NOBEL LECTURES: CHEMISTRY 377–432 (Tore Frängsmyr & Sture Forsén 1980).

sequenced.[42] Consequently, the total amount of known DNA sequence information forms an ever growing mountain of data.

E. Information

Long before the discovery of the DNA double helix, Archibold Garrod suggested that "genes work by controlling the synthesis of specific enzymes."[43] Employing *Neurospora* fungus as a model organism, George Beadle and Edward Tatum were able to confirm this "[O]ne [G]ene-[O]ne [E]nzyme [H]ypothesis" by demonstrating that "one gene controlled a single chemical reaction, which in turn was regulated by a specific enzyme."[44] This led to the parsimonious inference that "genetic information within genes determines the order of the 20 different amino acids within the polypeptide chains of proteins." [45]

Close on the heels of announcing the DNA double helix, Watson and Crick published the "Genetical Implications of the Structure of Deoxyribonucleic Acid," in which they postulated that the DNA sequence of a gene corresponded precisely to the amino acid sequence of a corresponding polypeptide.[46] In 1958, Crick elaborated on this hypothesis, proposing that "the specificity of a piece of nucleic acid is expressed solely by the sequence of its bases, and that this sequence is a (simple) code for the amino acid sequence of a particular protein."[47]

The first to crack part of the genetic code were Marshall Nirenberg and Heinrich Matthaei, who "observed in 1961 that the addition of the synthetic polynucleotide poly U (UUUUU . . .) to a cell-free system capable of making proteins leads to the synthesis of polypeptide chains containing only the amino acid phenylalanine. The nucleotide groups UUU thus must specify phenylalanine."[48] This discovery set off a race to break the rest of the genetic code, and "[b]y 1967 the code was essentially completed."[49] Understanding DNA as encoding information via a specific language of codons revealed a duality in the nature of genes and DNA: they were simultaneously physical and informational. The

---

42 *See, e.g.*, E. S. Lander et al., *Initial Sequencing and Analysis of the Human Genome*, 409 NATURE 860 (2001); J. Craig Venter et al., *The Sequence of the Human Genome*, 291 SCI. 1304 (2001).

43 WATSON ET AL., *supra* note 28, at 19.

44 LILY E. KAY, WHO WROTE THE BOOK OF LIFE 52 (2000).

45 WATSON ET AL., *supra* note 28, at 19.

46 J. D. Watson & F. H. C. Crick, *Genetical Implications of the Structure of Deoxyribonucleic Acid,* 171 NATURE 964, 964 (1953).

47 F. H. C. Crick, *On Protein Synthesis*, 12 SYMP. SOC'Y FOR EXPERIMENTAL BIOLOGY 138, 152 (1958) [hereinafter *On Protein Synthesis*].

48 WATSON ET AL., *supra* note 28, at 37.

49 KAY, *supra* note 44, at 330.

informational nature of DNA separates it qualitatively from other molecules. Consequently, "[t]he human genome is now generally viewed as an information system and, more specifically, as a 'Book of Life' written in the language of DNA, or DNA code, to be read and edited."[50]

F. Program

In his famous summation of genetics, Crick once remarked that "DNA makes RNA, RNA makes protein, and proteins make us."[51] Elaborating on this simple recipe, François Jacob and Jacques Monod suggested that "the genome contains not only a series of blue-prints, but a co-ordinated program of protein synthesis and the means of controlling its execution."[52] Jacob and Monod uncovered a complex network of interacting genetic elements that came to be known as the "operon" model of gene regulation.[53] This operon model suggested a "genetic program," comprising, as it did, "a linked cluster of regulatory elements and structural genes whose expression is coordinated by the product of a regulator gene situated elsewhere in the genome."[54]

James Bonner expanded on the concept of a genetic program in his 1965 book entitled *The Molecular Biology of Development*. Bonner deliberately drew a close analogy between computers and cells. He described a hierarchy of programs, including a "'master programme constituted in turn of a set of subprogrammes or subroutines,'"[55] with the latter further subdivided into "a list of cellular instructions or commands."[56]

Although organisms and their cells are not electronic computers, and genes are not written in software code, the similarities are striking. In fact, one of the major goals of synthetic biology is "to create a programmable microorganism from scratch."[57]

---

50 *Id.* at 1.

51 KELLER, *supra* note 21, at 54.

52 François Jacob & Jacques Monod, *Genetic Regulatory Mechanisms in the Synthesis of Proteins*, 3 J. MOLECULAR BIOLOGY 318, 354 (1961).

53 *See* LEWIN, *supra* note 15, at 858 ("An operon is a unit of bacterial gene expression and regulation, including structural genes and control elements in DNA recognized by regulator gene product(s).").

54 KAY, *supra* note 44, at 57.

55 *Id.* at 85, 134.

56 *Id.* at 85–86.

57 Arjun Bhutkar, *Synthetic Biology: Navigating the Challenges Ahead*, 8 J. BIOLAW & BUS. 19, 20 (2005).

III. PATENT PROTECTION FOR DNA

A. Origin

During 1973 and 1974, Stanley Cohen and Herbert Boyer invented a method for removing a specific fragment of DNA from one organism and introducing it into the genome of a different organism.[58] Significantly, they received a patent for their invention, which was entitled "Process for producing biologically functional molecular chimeras."[59] This method proved to be important to modern molecular biology and foundational to the biotechnology industry. Within several years, biologists had genetically engineered the eubacterium, *Escherichia coli*, successfully to express the gene encoding the human hormone somatostatin.[60] Of similar importance to the development of the biotechnology industry was a legal decision about a patent application filed by Dr. Ananda Chakrabarty in 1972. Rather than claim a method, this patent application involved a genetically modified organism. Among other inventions claimed in the patent application was a "human-made, genetically engineered bacterium . . . capable of breaking down multiple components of . . . oil."[61] In 1980, the United States Supreme Court upheld the patentability of "[a] bacterium from the genus *Pseudomonas* containing therein at least two stable energy-generating plasmids."[62] Not only did the Supreme Court appear to support the patentability of organisms, it also appeared to approve of the patentability of isolated DNA. Biotechnology began advancing rapidly on both scientific and legal fronts.

Complex biological molecules—such as DNA, RNA, and polypeptides—became the subject of patent claims beginning in the early 1970s. Previously, patents had successfully claimed methods that involved polypeptides and proteins, but in 1971, composition claims were issued for a polypeptide in U.S. Patent Number 3,607,370,[63] and for a protein in U.S. Patent Number 3,619,206.[64] Additionally, in 1972, a composition claim for a peptide was issued in U.S. Patent Number

---

58 Sally Smith Hughes, *Making Dollars Out of DNA—The First Major Patent in Biotechnology and the Commercialization of Molecular Biology, 1974–1980*, 92 ISIS 541, 541 (2001).

59 U.S. Patent No. 4,237,224 (filed Nov. 4, 1974).

60 Keiichi Itakura et al., *Expression in* Escherischia Coli *of a Chemically Synthesized Gene for the Hormone Somatostatin*, 198 SCI. 1056 (1977).

61 Diamond v. Chakrabarty, 447 U.S. 303, 305 (1980).

62 U.S. Patent No. 4,259,444 (filed June 7, 1972) (emphasis added).

63 Pressure-Sensitive Adhesive Tape Comprising Gluten Hydrolypate Derivatives, U.S. Patent No. 3,607,370 (filed May 29, 1969).

64 Modified Proteins, U.S. Patent No. 3,619,206 (filed May 21, 1969).

3,645,689.[65] A similar progression occurred for DNA and genes. A patent claim including "DNA" was issued in 1973,[66] the word "gene" occurred in an issued claim in U.S. Patent Number 3,710,511,[67] and 1978 saw the issuance of claims in U.S. Patent Number 4,116,770 that were directed to phenotypic traits encoded by specific genes.[68] However, it was not until two years after the decision in *Diamond v. Chakrabarty* that U.S. Patent Number 4,363,877 was issued with composition claims that included genes, in this case encoding growth hormone polypeptide.[69] These claims were the first to cover genes themselves,[70] which encoded "[h]uman chorionic somatomammotropin" and "the growth hormone of an animal species."[71]

After the *Diamond v. Chakrabarty* decision in 1980, patent applications, with a necessary lag time, issued patents claiming genes and DNA rose rapidly. Annual filings of patent applications rose from hundreds in 1984 to thousands from the mid-1990s to the early 2000s.[72] Annual patent issuances rose from hundreds in the late 1980s to thousands during the late 1990s until the late 2000s.[73]

B. Effect

Patent protection affords patent owners the right to exclude others from making, using, offering to sell, or selling patented genes within the United States. In addition, patent owners can prevent others from importing patented genes into the United States.[74] The availability of patent protection for DNA sequences has played a key role in attracting investment to the biotechnology industry. In fact, for an industry that has yet to produce a profit as a whole, patent portfolios covering genes may constitute one of the most valuable assets owned by biotechnology

---

65 Method and Apparatus for Analyzing Proteins, U.S. Patent No. 3,645,689 (filed Apr. 9, 1970).

66 Diagnostic Method Utilizing Synthetic Deoxyrilionucleotide Oligomer Template, U.S. Patent No. 3,755,086 (filed Feb. 9, 1971).

67 Procedures for Use of Genic Male Sterility in Production of Commercial Hybrid Maize, U.S. Patent No. 3,710,511 (filed Apr. 21, 1971).

68 Waxy Barley Starch with Unique Self-Liquifying Properties, U.S. Patent No. 4,116,770 (filed Feb. 27, 1975).

69 Recombinant DNA Transfer Vectors, U.S. Patent No. 4,363,877 (filed Apr. 19, 1978).

70 *Gene Patents and Global Competition Issues – Protection of Biotechnology Under Patent Law*, GENETIC ENGINEERING & BIOTECHNOLOGY NEWS (Jan. 1, 2006), http://www.genengnews.com/articles/chitem.aspx?aid=1163&chid=0.

71 Recombinant DNA Transfer Vectors, U.S. Patent No. 4,363,877 (filed Apr. 19, 1978).

72 Torrance, *Gene Concepts, Gene Talk, and Gene Patents*, *supra* note 1, at 157–91.

73 *Id.*

74 35 U.S.C. § 271(a) (2006).

companies.[75] Sheila Jasanoff has described various important ways in which patents have supported the development of biotechnology:

> Especially in the United States, patents played a foundational role in the development of the biotechnology industry at several levels. First, the extension of patents to the life sciences created new classes of property rights in things that were previously outside the realm of what could be owned, or even thought of as subject to ownership claims. As a result, these objects became commodities that could have value, be exchanged, circulate in markets, and foster productivity. Second, much of the early development of biotechnology occurred before there were any marketable products, and patents were the only evidence for eager venture capitalists that there might be something of future value to justify present investment. Third, patents provided some assurance to jittery investors that they would not be mired in endless legal wrangling if commercially useful products ever came on line. Fourth, patents proved to be a way of sorting out the competing claims of participants in an increasingly complex web of invention that linked together the disparate interests of patients, research subjects, farmers, academic researchers, universities, start-up firms, government, and industry.[76]

In short, patents appear to have played an especially crucial role in justifying the huge investments in research, development, and regulatory compliance that biotechnology and pharmaceutical companies face in discovering gene-based drugs and bringing them to market.[77]

---

[75] *See* John Golden, *Biotechnology, Technology Policy, and Patentability: Natural Products and Invention in the American System*, 50 EMORY L.J. 101 (2001).

[76] SHEILA JASANOFF, DESIGNS ON NATURE: SCIENCE AND DEMOCRACY IN EUROPE AND THE UNITED STATES 203–04 (2005).

[77] In their recent study of the role that the patent system plays in spurring innovation, James Bessen and Michael J. Meurer suggest the patent system may indeed promote innovation in the pharmaceutical/biotechnology industry. JAMES BESSEN & MICHAEL J. MEURER, PATENT FAILURE: HOW JUDGES, BUREAUCRATS, AND LAWYERS PUT INNOVATORS AT RISK 85–88 (2008). As Bessen and Meurer stated, "[t]he evidence certainly is consistent with the notion that patents encourage American pharmaceutical R&D." James Bessen & Michael J. Meurer, *Do Patents Stimulate R&D Investment and Promote Growth?*, PATENTLYO BLOG (Mar. 13, 2008), http://www.patentlyo.com/patent/2008/03/do-patents-stim.html.

C. Criticism

Gene patents are not without their critics. Controversy over patenting DNA is due, in large part, to the fact that even human genes have been considered eligible for patentable protection. According to a study by Fiona Murray and Kyle Jensen, approximately 20% of the known genes in the human genome have been claimed in patents issued by the United States Patent and Trademark Office ("USPTO").[78] The ethics of allowing this "'gold rush'" have been questioned,[79] as have the practical threats to genetic research allegedly posed by the consequent tragedy of the "[a]nticommons."[80] More sensational claims misinterpret the scope of the patent grant, which, in the case of DNA derived from genomic sources, usually extends only to "isolated" or "purified" DNA. Nevertheless, Devanand Crease and George Schlich have warned that, "[t]o the person in the street, the grant of a patent covering all potential uses of these genes raises the visceral fear of corporate interests claiming ownership over our very bodies!,"[81] while Michael Crichton raised this 2007 alarm in the New York Times: "YOU, or someone you love, may die because of a gene patent . . . . Gene patents are now used to halt research, prevent medical testing and keep vital information from you and your doctor."[82]

Despite the sincerity of these anxieties, the data appears to tell a different story. In a comprehensive analysis of all human gene patents identified by Murray and Jensen, Chris Holman found that "not one of the 4,270 patents in the dataset has ever been found to have been infringed or been the basis of a preliminary injunction."[83] Nevertheless, anxieties over the patenting of DNA appear to have struck a chord with judges and politicians, and the last several years have presented a decidedly more hostile atmosphere for DNA patents. Three specific events exemplify this shift: (1) *In re Fisher*, (2) the *Genomic Research and*

---

78 Kyle Jensen & Fiona Murray, *Intellectual Property Landscape of the Human Genome*, 310 SCI. 239 (2005).

79 Tom Hollon, *Gene Patent Revisions to Remove Some Controversies*, 6 NATURE MED. 362 (2000).

80 Michael A. Heller & Rebecca S. Eisenberg, *Can Patents Deter Innovation? The Anticommons in Biomedical Research*, 280 SCI. 698 (1998).

81 Devanand Crease & George Schlich, *Is There a Future for 'Speculative' Gene Patents in Europe?*, 2 NATURE REVS. DRUG DISCOVERY 407 (2003) (addressing patents claiming the human genes, BRCA1 and BRCA2, owned by Myriad Genetics, and used to diagnose propensity for developing breast cancer).

82 Michael Crichton, Op-Ed., *Patenting Life*, N.Y. TIMES, Feb. 13, 2007, at A2.

83 Christopher M. Holman, *The Impact of Human Gene Patents on Innovation & Access: A Survey of Human Gene Patent Litigation*, 76 U. MO. KAN. CITY L. REV. 295, 353–54 (2007) (emphasis omitted) (citations omitted).

Accessibility Act ("GRAA"), and (3) *Ass'n for Molecular Pathology v. U.S. Patent & Trademark Office*.

D. Response

In 2005, the Court of Appeals for the Federal Circuit heard an appeal regarding the patentability of fragments of genes called "expressed sequence tags" ("ESTs"), which are capable of identifying specific DNA sequences in maize genes.[84] The Federal Circuit decided that claims to these ESTs were invalid because they lacked utility and enablement. In explaining its decision, the court argued that:

> The basic quid pro quo contemplated by the Constitution and the Congress for granting a patent monopoly is the benefit derived by the public from an invention with substantial utility. Unless and until a process is refined and developed to this point—where specific benefit exists in currently available form—there is insufficient justification for permitting an applicant to engross what may prove to be a broad field.[85]

Rather than providing a mere benefit in conducting further research, the court suggested that a claimed invention should provide "some immediate benefit to the public" in order to be patentable.[86] In the wake of *In re Fisher*, there was significant uncertainty about what characteristics, if any, might make a DNA fragment eligible for patentability. This doubt spread to the patentability of longer stretches of DNA, including whole genes.

On February 7, 2007, two members of the House of Representatives, Xavier Becerra (Democrat of California) and Dave Weldon (Republican of Florida), introduced a bill that would "end[ ] the practice of gene patenting [including all] genetic material, naturally-occurring or modified."[87] The bill, the GRAA, would amend the Patent Act by adding a new section specifically addressing DNA patents. Proposed Section 106 would end the patentability of genes and other DNA sequences by providing that "[n]otwithstanding any other provision of law, no patent may be obtained for a nucleotide sequence, or its functions or correlations, or the naturally occurring products it specifies." Soon after its introduction in the House, the GRAA attracted the attention of the

---

84 *In re* Fisher, 421 F.3d 1365, 1367 (Fed. Cir. 2005).

85 *Id.* at 1371 (emphasis omitted).

86 *Id.* (emphasis omitted) (citation omitted).

87 153 CONG. REC. E315–16 (daily ed. Feb. 9, 2007) (statement of Rep. Xavier Becerra).

Subcommittee on Courts, the Internet and Intellectual Property. A hearing called "Stifling or Stimulating: The Role of Gene Patents in Research and Genetic Testing" was held on October 30, 2007. The outlook for passage of the GRAA is uncertain. It has not been passed by either the House or the Senate, and one of its sponsors, Representative Weldon, is no longer a member of the House.

However, the aims of the GRAA are strikingly congruent to the results of *Ass'n for Molecular Pathology v. U.S. Patent & Trademark Office*. In fact, the decision by Judge Sweet in *Ass'n for Molecular Pathology v. U.S. Patent & Trademark Office* appeared to echo Becerra's assessment of the patent requirement that DNA must be "isolated and purified" as "mere wordplay."[88] Sweet similarly described this requirement as a "lawyer's trick."[89] Furthermore, a proposed last-minute amendment to the Patent Reform Act of 2011 would have created Section 287(d) to create a safe harbor provision to limit infringement remedies against anyone who performs a genetic diagnostic test to give a patient a second opinion.[90] The American Civil Liberties Union ("ACLU") and its allies opposed this amendment, fearing that it "would fail to block all patent holder objections to such testing, fails to address the many other limitations on scientific research arising out of the issuance of such patents, and risks allowing gene patent holders to argue that Congress implicitly endorses the validity of such patents."[91] Although the amendment was subsequently withdrawn, it resurfaced in the patent reform bill that passed the House on June 23, 2011.[92] Obviously, there continues to be significant Congressional interest in curtailing patents that claim DNA sequences.

The most important event thus far in darkening the prospects for DNA patents began on May 12, 2009. The ACLU represented several female patients as well as a number of sympathetic organizations. Together, they sued the USPTO, a Utah-based biotechnology firm called Myriad Genetics, and the Directors of the University of Utah Research Foundation in the Federal court for the Southern District of New York.

---

88 *Id.*

89 Ass'n for Molecular Pathology v. Myriad Genetics, 702 F. Supp. 2d 181, 185 (S.D.N.Y. 2010).

90 Amendment to America Invents Act of 2011, H.R. 1249, 112th Cong. (1st Sess. 2011).

91 Letter from Laura W. Murphy, Dir., Washington Legislative Office, to Chairman David Dreier, Comm. on Rules, U.S. House of Representatives (June 15, 2011) (on file with the ACLU).

92 America Invents Act of 2011, H.R. 1249, 112th Cong. (1st Sess. 2011). This bill was subsequently enacted into law. Leahy-Smith America Invests Act. Pub. L. No. 112-29, 125 Stat. 284.

The cause of action in the lawsuit was opposition to the patenting of human genes and diagnostic uses thereof.[93] The complaint stated:

> Every person's body contains human genes, passed down to each individual from his or her parents. These genes determine, in part, the structure and function of every human body. This case challenges the legality and constitutionality of granting patents over this most basic element of every person's individuality. . . . [as well as granting patents covering] the concept of looking at or comparing human genes, and correlations found in nature between certain genes and an increased risk of breast and/or ovarian cancer. . . .[94]

In addition to attracting wide publicity, the ACLU lawsuit surprised many when it was taken seriously by the court.

The ACLU argued that the patent eligibility of genes posed a direct and serious threat to those susceptible to breast and ovarian cancers. Breast cancer afflicts about 13% of women in the United States over their lifetimes, is newly diagnosed in roughly 200,000 women annually, and is responsible for about 40,000 deaths per year, making it the third largest cause of cancer deaths.[95] The lifetime risk of developing ovarian cancer is about 1.7%. However, for carriers of the BRCA ("breast cancer") tumor suppressor gene mutations 1 ("BRCA1") and 2 ("BRCA2"), the lifetime probability of developing breast cancer and ovarian cancer rises dramatically to 36–85% and 20–60%, respectively.[96] Although knowing that one has BRCA1 or BRCA2 mutations is terrible news, it does allow the carrier and her physician to attempt to minimize other risks of developing cancer. As a result, Myriad Genetics has become a profitable company by acquiring patents covering both the BRCA1 and BRCA2 mutations and offering genetic tests to detect the presence of these mutations.[97] The ACLU argued that it is immoral to allow Myriad

---

93 Complaint at 1, 3, Ass'n for Molecular Pathology v. Myriad Genetics, 702 F. Supp. 2d 181 (S.D.N.Y 2010) (No. 2010-1406).

94 *Id.*

95 *Breast Cancer*, MAYO CLINIC (2011), http://www.mayoclinic.org/breast-cancer/.

96 NORTHWESTERN ASSOCIATION FOR BIOMEDICAL RESEARCH, APPENDIX A2, *available at* http://nwabr.org/sites/default/files/learn/bioinformatics/IntroAppendix.pdf (last visited Oct. 26, 2011); *see also* Matthew J. Piehl, *The Brave New World of Genetic Biobanks: International Lessons for a Potential United States Biobank*, 46 VAL. U. L. REV. 69, 93–94 (2011) (describing the harmful effects of the BRCA1 and BRCA2 gene mutations and a woman's right not to know she is carrying them).

97 *See* ENCYCLOPEDIA OF SCIENCE AND TECHNOLOGY COMMUNICATION 331 (Susanna Hornig Priest ed., 2010) ("The patents offer Myriad Genetics exclusive rights over the

Genetics to restrict access to such a beneficial diagnostic test by means of its patents. Myriad Genetics counter-argued that gene patents drove innovation into diagnostic tests, such as those it offered, by allowing companies a return on their investments in research, development, and regulatory approval.

The district court handed ACLU a decisive victory on March 29, 2010. On summary judgment, Judge Sweet held that neither genes nor genetic tests were eligible for patent protection. In a vigorously worded opinion, he undermined thirty years of DNA patent practice:

> The claims-in-suit directed to "isolated DNA" containing human BRCA1/2 gene sequences reflect the USPTO's practice of granting patents on DNA sequences so long as those sequences are claimed in the form of "isolated DNA." This practice is premised on the view that DNA should be treated no differently from any other chemical compound, and that its purification from the body, using well-known techniques, renders it patentable by transforming it into something distinctly different in character. Many, however, including scientists in the field of molecular biology and genomics, have considered this practice a "lawyer's trick" that circumvents the prohibitions on the direct patenting of DNA in our bodies but which, in practice, reaches the same result . . . . It is concluded that DNA's existence in an "isolated" form alters neither this fundamental quality of DNA as it exists in the body nor the information it encodes. Therefore, the patents at issue directed to "isolated DNA" containing sequences found in nature are unsustainable as a matter of law and are deemed unpatentable subject matter under 35 USC § 101.[98]

Myriad Genetics promptly appealed this summary judgment decision to the Federal Circuit. Fearing massive losses if their vast portfolios of DNA patents were also found invalid, the pharmaceutical and biotechnology industries strongly supported the appeal.

BRCA1 and BRCA2 genes and prevent others from further studying these genes without getting a license and paying royalties. Myriad established a monopoly for the genetic test, and currently there is no other way to test for the presence of BRCA mutations without infringing the Myriad patent.").

[98] Ass'n for Molecular Pathology v. Myriad Genetics, 702 F. Supp. 2d 181, 185 (S.D.N.Y. 2010).

Before the Federal Circuit ruled on the appeal, one of the original defendants, the United States (representing the USPTO), surprised supporters and opponents of Myriad Genetics alike when it changed sides to support, at least partially, positions of the ACLU. On October 29, 2010, the Department of Justice filed an *amicus curiae* brief on behalf of the federal government—including erstwhile defendant, the USPTO—arguing that "isolated but otherwise unaltered" human genes constitute unpatentable subject matter because they are products of nature under 35 U.S.C. § 101.[99] On July 29, 2011, a panel of three Federal Circuit judges reversed much of Judge Sweet's decision, and, in doing so, reaffirmed the eligibility of isolated DNA sequences for patent protection.[100] Notwithstanding this latest change of fortunes for gene patents, it is likely that the losing party, the Association for Molecular Pathology, will appeal this decision to the United States Supreme Court. If the Supreme Court does grant a writ of certiorari in this case, perhaps pairing it with another case, *Prometheus v. Mayo*,[101] in which it already granted certiorari to consider the patent-eligibility of biotechnology inventions, and given the new official position of the federal government, a decision finding at least some categories of DNA ineligible for patenting is possible.

The results of the recent trend to limit patents claiming DNA will not be known until two things happen: the federal courts finally decide the issues in *Ass'n for Molecular Pathology v. Myriad Genetics* and *Prometheus v. Mayo*, and Congress finally decides on what, if any, statutory reforms it supports. In the meantime, much uncertainty hovers over the future viability, vitality, and value of DNA patents. Copyright protection for DNA sequences offers an alternative.

IV. SYNTHETIC BIOLOGY

In its most optimistic conception, the field of synthetic biology promises nothing short of a brave new world. As Michael Specter observed in 2009, "[i]f the science truly succeeds, it will make it possible to supplant the world created by Darwinian evolution with one created by us."[102] However, the field has more modest immediate goals: "By

---

[99] Brief for the United States as Amicus Curiae in Support of Neither Party, Ass'n for Molecular Pathology v. Myriad Genetics, 702 F. Supp. 2d 181 (S.D.N.Y. 2010) (No. 2010-1406).

[100] Ass'n for Molecular Pathology v. U.S. Patent and Trademark Office, 653 F.3d 1329 (Fed. Cir. 2011).

[101] Prometheus Labs., Inc. v. Mayo Collaborative Servs., 628 F.3d 1347 (Fed. Cir. 2010).

[102] Michael Specter, *A Life of Its Own: Where Will Synthetic Biology Lead Us?*, THE NEW YORKER, Sept. 28, 2009, at 57.

using gene-sequence information and synthetic DNA, they are attempting to reconfigure the metabolic pathways of cells to perform entirely new functions, such as manufacturing chemicals and drugs. *Eventually*, they intend to construct genes—and new forms of life—from scratch."[103] The success of synthetic biology will largely depend on two fundamental technologies whose very names suggest the relevance of copyright: (1) reading DNA (via rapid and inexpensive sequencing); and (2) writing DNA (via rapid and inexpensive synthesis). To the long evolution of DNA concepts, from factor to particle to sequence to information to program, synthetic biology has the added possibility of viewing DNA as lego-like "BioBricks" that can be arranged and rearranged at will to build new structures and functions. In addition, synthetic biology has introduced the idea of biologists as creative authors engaged in literally writing the future of the book of life nucleotide by nucleotide.

A. Recombining

The methods of recombinant DNA developed by Cohen and Boyer demonstrated that DNA from a foreign source organism could be reliably spliced into the genome of a distinct host organism. They used restriction endonucleases to create a gap in a eubacterial plasmid, inserted foreign DNA into the gap, and used DNA ligase to splice the plasmid and foreign DNA together.[104] By showing that DNA could be recombined into arrangements not found in existing genomes,[105] Cohen and Boyer opened the door to the deliberate design of new genomes. The subsequent successful insertion of the human somatostatin gene into a eubacterial genome amplified this possibility by demonstrating the intercompatibility of DNA from very disparate phylogenetic sources.[106] These experiments and techniques heralded "the new era of 'synthetic biology' where not only existing genes are described and analyzed but also new gene arrangements can be constructed and evaluated."[107]

B. Programming

Biologists have long used the metaphor of the computer program to describe the function of DNA. Crick's succinct formulation anticipated

---

103 *Id.* at 56 (emphasis added).

104 Stanley N. Cohen et al., *Construction of Biologically Functional Bacterial Plasmids In Vitro*, 70 PROC. NAT'L ACAD. SCI. 3240, 3240 (1973).

105 *Id.* at 3244.

106 *E.g.*, Itakura et al., *supra* note 60, at 1056.

107 Waclaw Szybalski & Ann Skalka, *Nobel Prizes and Restriction Enzymes*, 4 GENE 181, 181–82 (1978).

subsequent, more explicit, uses of the metaphor: "DNA makes RNA, RNA makes protein, and proteins make us."[108] Jacob and Monod offered a more direct formulation, suggesting that "the genome contains not only a series of blue-prints, but a co-ordinated program of protein synthesis and the means of controlling its execution."[109] Bonner completed the progression by describing the genome as a "master programme constituted in turn of a set of subprogrammes or subroutines."[110] To be accurate, these formulations of DNA as computer programs depend on cells or organisms capable of reliably processing and implementing the instructions of their DNA programs. While natural organisms that have resulted from organic evolution may present challenges in this regard, a goal of synthetic biology is to introduce reliability and predictability through deliberate and careful engineering. To this end, Arjun Bhutkar asserts that "[a] primary objective of this nascent research area is to create a programmable microorganism from scratch."[111]

C. Engineering

What is currently called synthetic biology was once known as biological engineering. In the 1930s, MIT appointed Professor Joseph Warren Horton the inaugural head of "the newly created Department of Biological Engineering."[112] Despite the farsightedness of its founders, the Department changed its status at MIT several times. In 1998, the successor to the Department became a division of the MIT School of Engineering, and in 2005, it reemerged as an independent department.[113] In his 1958 acceptance speech of the Nobel Prize for Medicine, Edward L. Tatum offered his vision of how biology might transform itself into biological engineering:

> With a more complete understanding of the functioning and regulation of gene activity in development and differentiation, these processes may be more efficiently controlled and regulated, not only to avoid structural or

---

108 *On Protein Synthesis*, *supra* note 47, at 139.

109 Jacob & Monod, *supra* note 52, at 354.

110 JAMES BONNER, THE MOLECULAR BIOLOGY OF DEVELOPMENT 134 (1965).

111 Bhutkar, *supra* note 57, at 19, 20.

112 Gerald L. Zeitlin, *Professor Joseph Warren Horton (1889–1967): Biological Engineer*, 13 J. MED. BIOGRAPHY, 39, 39 (2005).

113 *Massachusetts Institute of Technology School of Engineering*, WIKIPEDIA, http://en.wikipedia.org/wiki/Massachusetts_Institute_of_Technology_School_of_Engineering (last modified July 2, 2011).

> metabolic errors in the developing organism, but also to produce better organisms.
> . . . [Understanding the genetic code] may permit the improvement of all living organisms by processes which we might call biological engineering.[114]

In 2005, Drew Endy published a series of suggestions entitled Foundations for Engineering Biology, which outlined how biology could finally achieve its promise as an engineering discipline.[115] He proposed three general principles: (1) standardization; (2) decoupling; and (3) abstraction.[116] Standardization required "the definition, description and characterization of the basic biological parts, as well as standard conditions that support the use of parts in combination and overall system operation."[117] Decoupling would break larger problems into a set of discrete and smaller problems that could be solved separately. Then, once its constituent smaller problems had been solved, a larger problem could be solved.[118] Abstraction would involve separating biological engineering problems into hierarchical levels of complexity ("abstraction hierarchies") and then reengineering basic biological structures and functions into simpler components.[119] Biological engineering would produce standard biological parts capable of being combined into more complex biological devices, which in turn could be combined into even more complex biological systems.[120]

Work has already begun on producing biological parts. The Registry of Standard Biological Parts "is a continuously growing collection of genetic parts that can be mixed and matched to build synthetic biology devices and systems."[121] As of June 22, 2011, the Registry contained 15,177 genetic parts consisting of deposited DNA sequences.[122] Furthermore, the International Open Facility Advancing Biotechnology ("Biofab") was founded in December 2009 "as the world's first biological design-build facility."[123] Funded by the National Science Foundation,

---

114 Edward Tatum, *A Case History in Biological Research,* NOBELPRIZE.ORG (June 30, 2011), http://nobelprize.org/nobel_prizes/medicine/laureates/1958/tatum-lecture.html.

115 DREW ENDY, FOUNDATIONS FOR ENGINEERING BIOLOGY 449–53 (2005).

116 *Id.* at 450–52.

117 *Id.* at 450.

118 *Id.* at 451.

119 *Id.* at 451–52.

120 *Id.*

121 REGISTRY OF STANDARD BIOLOGICAL PARTS, http://partsregistry.org/Main_Page (last visited June 30, 2011).

122 *Statistics Snapshot,* REGISTRY OF STANDARD BIOLOGICAL PARTS, http://partsregistry.org/cgi/partsdb/Statistics.cgi (last visited June 30, 2011).

123 *About the Biofab,* BIOFAB, http://www.biofab.org/about (last visited June 30, 2011).

"[o]nce fully operational the Biofab facility will be capable of producing tens of thousands of professionally engineered, high-quality standard biological parts each year."[124] Whether called biological engineering or synthetic biology, a new kind of biology in which DNA is deliberately designed, written, and authored has arrived.

V. COPYRIGHT PROTECTION FOR DNA[125]

Works of genetic authorship fit within the existing framework of copyright law. Congress does not need to amend the Copyright Act itself; rather, courts must recognize that DNA is already copyrightable subject matter. In part, this testifies to the flexibility of the Copyright Act itself. However, DNA is also a molecule with properties that uniquely preadapt it to eligibility for copyright protection. It is composed of an alphabet of nucleotides, A, G, T, and C, whose specific order creates meaning in a polynucleotide sequence.[126] The meaning that resides in the order of nucleotides is information that is easily decipherable both to the cells that harbor DNA and to human readers of nucleotide sequences.

Sequences of DNA have yet to be widely recognized as eligible for copyright protection. This necessitates consideration of several aspects of copyright law that might seem to be hurdles for DNA. These include the requirements of statutory subject matter, originality, authorship, and expression. In addition, there are necessary review bars to copyright eligibility, such as functionality. Some of the information in DNA is indeed functional. Nevertheless, due to redundancy in the genetic code and to stretches of apparently "junk" DNA, much opportunity exists for meaning—and expression—outside of the context of function. The case for DNA copyright, despite functionality, is bolstered by the eligibility of computer software for copyright protection because of similarities between DNA sequences and computer algorithms, such as their ability to encode functions. Sequences of DNA, especially synthetic DNA,

---

124 *Id.*

125 *See infra* Part IV (expanding upon the discussion of DNA Copyright in Andrew W. Torrance, *Synthesizing Law for Synthetic Biology*, *supra* note 1, at 642–48).

126 RNA is a biological molecule very similar to DNA. Instead of A, C, G, and T, RNA is composed of A, C, G, and uracil ("U"). Although RNA sometimes forms the genome of certain viruses, it more commonly acts as a messenger that carries information from DNA for the synthesis of polypeptides (that is, messenger RNA ("mRNA")), carries amino acids to the site of polypeptide synthesis (transfer RNA ("tRNA")), forms ribosomes (ribosomal RNA ("rRNA")), or resides in the cell nucleus (small nuclear RNA ("snRNA")). The case for copyright protection of RNA and amino acid sequences is very similar to that for DNA. In the absence of specific differences, this article will use "DNA" as a shorthand for DNA, RNA, and polypeptides.

already fit within the eligibility requirements of the Copyright Act. Recognition of this is the next step.

If recognized, copyright protection may offer an alternative to patent protection. For example, copyright protection is capable of producing a socially desirable balance of restricted and permissible uses of DNA sequences. It achieves this balance by replacing the strict liability regime of patent law with the more flexible fair use defense and the fostering of a feasible open source regime.

A. History

A number of authors have already discussed the applicability of copyright law to DNA sequences.[127] Irving Kayton was the first to address the issue in 1982. Initially, he assumed that DNA was uncopyrightable. As he described, "every intellectual and emotional prejudice, both sophisticated and primitive, to which he is subject opposed coming to the conclusions finally reached. Copyright protection for engineered DNA sequences seemed ludicrous."[128] Yet, careful analysis of the Copyright Act changed his mind decisively. Kayton summarized his conclusions as follows:

> virtually all original works of a genetic scientist are copyrighted automatically when he creates them; the scientist generally can enforce his copyrights; those copyrights may provide more effective protection than other forms of intellectual property in many circumstances; and copyright protection for genetically engineered works appears within the constitutional limits on Congressional power.[129]

Writing later in the decade, and with the benefit of a fuller flowering of the biotechnology industry, Dan Burk also came to the conclusion that DNA constituted subject matter eligible for copyright protection. Burk suggested that copyright protection could extend "to encompass both sequences of nucleotide bases and their written representation."[130]

---

127 *E.g.*, Burk, *supra* note 7, at 469; Tani Chen, *Can a Biological Sequence Be Copyrighted?*, INTELL. PROP. & TECH. L.J., Mar. 2007, at 1; Duncan M. Davidson, *Common Law, Uncommon Software*, 47 U. PITT. L. REV. 1037, 1104–05 (1986); Kayton, *supra* note 6, at 191; Smith, *supra* note 7, at 1096–1108; James G. Silva, *Copyright Protection of Biotechnology Works: Into the Dustbin of History?*, B.C. INTELL. PROP. & TECH. F. (Jan. 28, 2000), http://www.bc.edu/bc_org/avp/law/st_org/iptf/articles/content/2000012801.html.

128 Kayton, *supra* note 6, at 218.

129 *Id.* at 192.

130 Burk, *supra* note 7, at 496.

Although not all analyses have supported the desirability of copyright protection for DNA, none of the authors to consider this issue have refuted its eligibility under the Copyright Act. Rather, these analyses have tended to reject DNA copyrightability on grounds that either: (1) the DNA sequences in question were natural in origin, and, thus, lacked proper authorship or originality; or (2) public policy considerations, rather than existing copyright law, militated against protection. Furthermore, these previous analyses have not had the opportunity to consider the impact of the recent field of synthetic biology, a field based on the *de novo* design and construction of DNA under the direction of human creativity. Where non-synthesized DNA may often be copyright eligible, the arguments for copyright protection of synthetic DNA sequences are *a fortiori*.

B. Requirements

Copyright protection applies to "original works of authorship fixed in any tangible medium of expression, now known or later developed, from which they can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device."[131] Fixation can occur in any "form, manner, or medium."[132] However, the mode of fixation must be "sufficiently permanent or stable to permit it to be perceived, reproduced, or otherwise communicated for a period of more than transitory duration."[133] Since DNA is composed of stable chemical nucleotides, DNA sequences should easily meet this requirement. Furthermore, DNA possesses definite sequences of nucleotides that can easily be determined,[134] copies of DNA may be synthesized routinely and in effectively unlimited quantities,[135] and molecular DNA has been known to last for at least many thousands of years with its nucleotide sequence intact.[136] The authorship requirement might pose a barrier to the copyrightability of genes and other DNA sequences derived entirely from natural genomes. A challenge would be posed by 17 U.S.C. § 102, which provides that "[c]opyright protection subsists . . . in original works of authorship."[137] By analogy, someone

131 17 U.S.C. § 102 (2006).
132 H.R. REP. NO. 94-1476, at 52 (1976).
133 17 U.S.C. § 101 (2006).
134 *See, e.g.*, F. Sanger et al., *DNA Sequencing with Chain-Terminating Inhibitors*, 74 PROC. NAT'L ACAD. SCI. U.S.A. 5463 (1977).
135 *See, e.g.*, 2 JOSEPH SAMBROOK & DAVID W. RUSSELL, MOLECULAR CLONING: A LABORATORY MANUAL 8.4 (3d ed. 2001).
136 *See, e.g.*, Eske Willerslev & Alan Cooper, *Ancient DNA*, 272 PROC. ROYAL SOC'Y B. 3, 3–5 (2005).
137 17 U.S.C. § 102 (2006).

other than the author could not claim copyright protection for a preexisting manuscript simply by discovering its existence.[138] However, synthetic biology can involve the design and construction of new, human-designed DNA sequences. Here the synthetic biologist designs the particular DNA sequence and "writes" it when she synthesizes it.[139] Since there is an author in this case, such DNA sequences should qualify as "original works of authorship." Furthermore, although DNA sequences lack the explicit statutory recognition as copyrightable subject matter that computer software possesses, synthetic DNA sequences should be eligible for copyright protection under the expansive interpretation of "works of authorship" manifested by Congress in the legislative history of the 1976 amendments to the Copyright Act.[140] Finally, DNA sequences can be "perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device."[141] The genetic code of DNA is well understood by biologists, and DNA sequences are easily reproduced.[142] Furthermore, machines and routine laboratory methods allow the specific nucleotides in DNA sequences to be determined.[143]

Originality is another requirement of copyrightability. All sequences of DNA are composed of existing nucleotides, each of which is individually not new. Some DNA sequences are "recombinant" assemblages of existing nucleotide sequences ligated together. If the focus of analysis were individual nucleotides or constituent sequences, originality of DNA might be in question. However, in *Roth Greeting Cards v. United Greeting Cards Co.*, an analogous case of greeting cards

---

138 Of course, non-authors may obtain copyright protection through contractual means for works authored by others.

139 In fact, fixing a DNA sequence via more conventional tangible forms of expression, such as writing the nucleotide sequence down on paper, may also suffice.

140 *See* H.R. REP. NO. 94-1476, at 51 (1976) (describing the history of expansively interpreting "works of authorship" to include new and varied forms of information).

> The history of copyright law has been one of gradual expansion in the types of works accorded protection, and the subject matter affected by this expansion has fallen into two general categories. In the first, scientific discoveries and technological developments have made possible new forms of creative expression that never existed before. In some of these cases the new expressive forms—electronic music, filmstrips, and computer programs, for example—could be regarded as an extension of copyrightable subject matter Congress had already intended to protect, and were thus considered copyrightable from the outset without the need of new legislation.

Id.

141 17 U.S.C. § 102.

142 *See, e.g.*, SAMBROOK & RUSSELL, *supra* note 135, at 8.4–8.17.

143 *See, e.g.*, Sanger et al., *supra* note 134, at 5463.

whose constituent parts were not new, the court pointed out that the "proper analysis of the problem requires that all elements of each [work] . . . be considered as a whole. . . . Considering all of these elements together, the Roth cards are, in our opinion, both original and copyrightable."[144] Like greeting cards, recombinant or fully-synthetic DNA sequences should be considered as a whole. Under this perspective, recombinant or synthetic DNA sequences are also likely to qualify as original. Furthermore, novel nucleotides can also be used to make nucleotide sequences. In 2011, geneticists Farren Isaacs, George Church, and others developed a method of creating organisms with genetic codes different from those of existing organisms.[145] Thus, not only can synthetic biology create DNA sequences never before seen outside the laboratory, entirely new genetic codes can be developed as alternatives to existing codes based on DNA or RNA.

C. Subject Matter

There is no explicit mention of DNA sequences in 17 U.S.C. § 102, nor do any of the eight enumerated categories of copyrightable subject matter explicitly include DNA sequences. There are, however, several significant respects in which DNA, genes, arrays of genes, and genomes (not to mention their RNA and polypeptide products) fit within the "literary works" category,[146] both generally and as computer programs. Like the English alphabet of twenty-six letters, DNA is composed of an alphabet of four nucleotide "letters": A, T, G, and C.[147] Triplets of these nucleotide letters form "codons" that correspond to specific amino acids. When strung together in a linear chain, amino acids comprise polypeptides. A synthetic biologist can "write" strings of nucleotides (for example, genes) in any pattern she wishes. Some patterns of nucleotide letters could be written to produce specifically desired linear chains of amino acids. At a higher level of organization, a synthetic biologist could compose arrays of multiple synthetic genes in particular patterns to produce complex results inside and outside of cells. Literary works are defined in § 101 as "works . . . expressed in words, numbers,

---

144 429 F.2d 1106, 1109 (9th Cir. 1970). This case relied on the Copyright Act of 1909 and was therefore superseded by the new Copyright Act. *See* Loree Rodkin Mgmt. Corp. v. Ross-Simons, Inc., 315 F. Supp. 2d 1053, 1055 n.1 (C.D. Cal. 2004) (explaining that *Ross* did not control that case because *Ross* relied on the 1909 Copyright Act which was amended by the Copyright Act of 1976).

145 Farren J. Issacs et al., *Precise Manipulation of Chromosomes in Vivo Enables Genome-Wide Codon Replacement*, 333 SCI. 348, 348–353 (2011).

146 17 U.S.C. § 102 (2006).

147 A similar molecule, RNA, is composed of adenine, uracil (instead of thymine), guanine, and cytosine. The RNA alphabet is A, U, G, and C.

or other verbal or numerical symbols or indicia, regardless of the nature of the material objects . . . in which they are embodied."[148] Nucleotides, DNA, RNA, genes, amino acids, polypeptides, and proteins are certainly "indicia," and the letters used to denote nucleotides and amino acids, as well as the codes used to denote genes may also qualify as "verbal . . . symbols."[149] Furthermore, the statement "regardless of the nature of the material objects . . . in which they are embodied" could certainly include DNA or its related molecules.[150]

Eligibility for copyright protection is not restricted to the seven enumerated categories under § 102. Rather, the section introduces the enumerated categories with the phrase "include[s] the following categories."[151] In the "Definitions" section of the Copyright Act, § 101 explains that the "term[ ] 'including' . . . [is] illustrative and not limitative."[152] The House Report accompanying the 1976 Copyright Act reinforces this broad interpretation:

> The use of the word "include," as defined in [§] 101, makes clear that the listing is "illustrative and not limitative," and that the seven categories do not necessarily exhaust the scope of "original works of authorship" that the bill is intended to protect. Rather, the list sets out the general area of copyrightable subject matter, but with sufficient flexibility to free the courts from rigid or outmoded concepts of the scope of particular categories.[153]

When considered in conjunction with the expansive phrase in § 102, "any tangible medium of expression, now known or later developed,"[154] synthetic DNA sequences fit comfortably within the category of "literary works."[155]

D. Software

In 1974, the National Commission on New Technological Uses of Copyrighted Works ("CONTU") issued a report concluding that "computer programs, to the extent that they embody an author's original

---

148 17 U.S.C. § 101 (2006).
149 *Id.*
150 *Id.*
151 17 U.S.C. § 102.
152 17 U.S.C. § 101.
153 *See* H.R. REP. NO. 94-1476, at 53 (1976).
154 17 U.S.C. § 102.
155 *Id.*

creation, are proper subject matter of copyright."[156] The CONTU was careful to distinguish copyrightable subject matter, such as creative expression in computer software, from uncopyrightable subject matter, such as "idea[s], procedure[s], process[es], system[s], method[s] of operation, concept[s], principle[s], or discover[ies]."[157] Moreover, it emphasized that "one is always free to make the machine do the same thing as it would if it had the copyrighted work placed in it, but only by one's own creative effort rather than by piracy."[158] Formal recognition of computer software as copyrightable subject matter occurred in 1980, when Title 17 (the "Copyright Act") was amended to include explicit copyright protection for computer software.[159] Section 101 of the Copyright Act defines "computer program" as "a set of statements or instructions to be used directly or indirectly in a computer in order to bring about a certain result."[160] Although there are some special limitations on the exclusive rights conferred to owners of copyrights on computer software,[161] this form of expression is now routinely protected by copyright.

Synthetic biology is largely based on a conception of genes, cells, and organisms as programmable. In a measured version of this conception, Endy has suggested that "synthetic biology provides an opportunity to test the hypothesis that the genomes encoding natural biological systems can be 're-written,' producing engineered surrogates that might usefully supplant some natural biological systems."[162] However, as a more ambitious articulation has portrayed it, "[a] primary objective of [synthetic biology] is to create a programmable microorganism from scratch," [163] and it is increasingly possible to "program living organisms in the same way a computer scientist can program a computer."[164] Consequently, if computer software is copyrightable, perhaps "biological software" is, or ought to be, as well.

It is relatively easy for a human mind to understand the "meaning" of a DNA sequence. Once a proper reading frame has been determined

---

156 NAT'L COMM'N ON NEW TECHNOLOGICAL USES OF COPYRIGHTED WORKS, FINAL REPORT OF THE NATIONAL COMMISSION ON NEW TECHNOLOGICAL USES OF COPYRIGHTED WORKS 1, 2 (1978) [hereinafter NAT'L COMM'N].

157 *Id.* at 18 (citing 17 U.S.C. § 102 (b)).

158 *Id.* at 21.

159 *See* H.R. REP. NO. 96-1307, pt. 1, at 23–24 (1980).

160 17 U.S.C. § 101 (2006).

161 17 U.S.C. § 117 (2006).

162 ENDY, *supra* note 115, at 449.

163 Bhutkar, *supra* note 57, at 20.

164 *D.I.Y. Organisms,* ECOPOLIS, http://www.ecopolis.org/diy-organisms/ (last visited July 17, 2011).

for the sequence,[165] one only has to recognize triplets of nucleotides and assign corresponding amino acids to each triplet. Thus, someone of modest skill in genetics could examine a DNA sequence of 300 coding nucleotides, in proper reading frame, and then determine the specific 100 amino acid sequence of its corresponding polypeptide. By contrast, it is much more difficult for one of similar skill in computer software to understand the "meaning" of either object code or source code. With respect to computer software, both source code and object code are eligible for copyright protection.[166] Source code is a form of a computer program expressed in a programming language understandable to humans. Object code, by contrast, is a form of a computer program expressed in binary (that is, "1s" and "0s"); object code cannot generally be understood by the human mind. If object code is eligible for copyright protection, then, *a fortiori*, so should DNA sequences because they can be relatively easily understood.

Rather than portray DNA sequences as analogous to computer software, a synthetic biologist (and copyright law) might actually consider DNA sequences to be a form of computer software. A gene is a set of instructions for producing a polypeptide.[167] A cell (or even an organism), via the molecules, metabolic pathways, and signaling pathways it contains, acts in response to the set of instructions encoded in its genes to carry out a certain result. Thus, "a [gene encodes a] set of statements or instructions to be used directly or indirectly in a [cellular] computer in order to bring about a certain [metabolic or signaling] result."[168] Given that one of the primary goals of synthetic biology is to engineer cells and genes to become ever more like computers and computer software, as synthetic biology succeeds in making DNA appear more similar to computer software, DNA sequences will likely move towards copyrightability by analogy to computer software.

---

[165] LEWIN, *supra* note 15, at 860.

> A reading frame is one of three possible ways of reading a nucleotide sequence. Each reading frame divides the sequence into a series of successive triplets. There are three possible reading frames in any sequence, depending on the starting point. If the first frame starts at position 1, the second frame starts at position 2, and the third frame starts at position 3.

Id.

[166] *E.g.*, Apple Computer, Inc. v. Franklin Computer Corp., 714 F.2d 1240, 1248 (3d Cir. 1983).

[167] LEWIN, *supra* note 15, at 852. "A gene is the segment of DNA specifying a polypeptide chain." *Id.*

[168] *Cf.* 17 U.S.C. § 101 (2006) (showing how material about DNA sequences can fit into the existing definition of computer software). The bracketed material is added to make this point.

Alternatively, if cells and organisms are already computers, and genes are already software, then DNA sequences are already eligible for copyright protection.

Whether or not cells are computers and genes are computer software is largely an empirical question. Endy offers a number of examples, including:

> [a] DNA sequence that programmes a biofilm to take a photograph and perform distributed edge-detection on the light-encoded image . . . [a] DNA sequence that programmes any mammalian cell to count up to 256 in response to a generic input signal . . . [and a] DNA sequence that programmes any prokaryote to produce 25 gl^{-1} artemisinic acid.[169]

However, rather than characterizing any of these examples as science fiction or hopeful thinking, Endy notes that "each application is physically plausible, or is the direct extension of an already demonstrated result."[170] This suggests that synthetic biology is well on the way towards cells as computers and genes as computer software. The consequences for the copyrightability of synthetic DNA sequences are significant.

E. Functionality

Copyright law limits protection to works of authorship that do not monopolize a particular function.[171] If a DNA sequence of a synthetic gene were to represent the only way of producing an RNA or polypeptide with a particular function, then that sequence would not likely possess strong copyright protection. However, if multiple DNA sequences could produce the same RNA or polypeptide with a particular function, then any one individual sequence would likely have much stronger copyright protection. In addition, as long as a work of authorship is original, it cannot infringe the copyright of another work of authorship, even if the two works of authorship are identical. Thus, even a copyright protecting a particular synthetic DNA sequence would not prevent others from independently designing an identical or similar DNA sequence. As a consequence, independent invention of identical or similar synthetic DNA sequences would act as a counterbalance to any

---

169 ENDY, *supra* note 115, at 449.

170 *Id.*

171 *See* 2 MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT § 2.01[A] (Matthew Bender, rev. ed., 2009).

monopoly rights conferred on the first author. Copying would still constitute copyright infringement, but independent invention would be permissible. This would stand in stark contrast to the rights conferred by patents claiming DNA sequences because the strict liability regime of patent law does not relieve independent inventors from liability.

Patent law offers protection for functional creations. In fact, patent law includes an explicit requirement that an invention possess utility in order to be eligible for protection with a utility patent.[172] By contrast, courts have often hesitated to confer copyright protection to utilitarian works. In *Baker v. Selden*, the Supreme Court refused to endorse copyright protection for an accounting system explained in an otherwise copyrightable book entitled *Selden's Condensed Ledger of Bookkeeping Simplified*.[173] In explaining why the blank forms in the book were ineligible for copyright protection, the court stated that, "in using the [accounting system], the ruled lines and headings of accounts [on the blank forms] must necessarily be used as incident to it."[174] In other words, a work whose form is dictated solely by function is uncopyrightable. However, the mere fact that a work possesses functionality does not preclude it from copyright eligibility. To illustrate this point, the book at issue in *Baker v. Seldon* was copyrightable. As long as a work possesses adequate expression, it is eligible for copyright protection—even if it possesses functionality. Thus, DNA molecules are copyrightable to the extent their nucleotide sequences are not dictated by function. If the expression of an idea in an otherwise copyright-eligible work is entirely determined by functional considerations, copyright protection is not appropriate because expression and idea may have impermissibly merged. Some of the information in DNA is indeed functional, but due to redundancy in the genetic code and to stretches of apparently "junk" DNA, much opportunity exists for meaning—and expression—outside of the context of function. This may be especially true for synthetic DNA sequences. Synthesized strands of DNA may be deliberately designed to lack function.[175] They may be designed with

---

172 35 U.S.C. § 101 (2006).

173 Baker v. Selden, 101 U.S. 99, 100–01 (1880).

174 *Id.* at 104.

175 Press Release, J. Craig Venter Inst., First Self-Replicating Synthetic Bacterial Cell (May 20, 2010), *available at* http://www.jcvi.org/cms/press/press-releases/fulltext/article/first-self-replicating-synthetic-bacterial-cell-constructed-by-j-craig-venterinstitute-researcher/. For commentary of Venter's breakthrough, see Daniel G. Gibson et al., *Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome*, 329 SCI. 52 (2010); Jonathan Khan, *Synthetic Hype: A Skeptical View of the Promise of Synthetic Biology*, 45 VAL. U. L. REV. 1343, 1343 (2011); Kristine S. Knaplund, *Synthetic Cells, Synthetic Life, and Inheritance*, 45 VAL. U. L. REV. 1361, 1362–63, 1366 (2011); Stephen M. Maurer, *End of the Beginning or Beginning of the End? Synthetic Biology's Stalled Security Agenda and the Prospects for Restarting It*, 45 VAL. U.

function in mind, but may be expressed in many different permutations of specific nucleotides and corresponding amino acids. Additionally, some of their nucleotide positions (e.g., the first codon position) may be fixed for reasons of function, while other nucleotides (e.g., the second and third codon positions) are largely unfettered from functional constraints. Furthermore, even functionality is not an absolute bar to copyrightability for DNA sequences or other works. In fact, the functionality threshold a nucleotide sequence must exceed to be eligible for copyright protection is not high. Although some DNA molecules whose sequences are strictly functionally constrained may not surmount this threshold, many other less-constrained DNA molecules will not be precluded from copyrightability on grounds of functionality.

In practical terms, very short sequences of DNA encoding very short polypeptides would probably be uncopyrightable due to the very limited number of sequences capable of encoding the corresponding amino acids. Such sequences would likely lack sufficient expression for copyrightability. Such a limitation on copyrightability would ensure that short building-block sequences of DNA remained in the public domain. To illustrate, this might cast doubt on the strength of copyright protection that Illumina can expect for its oligonucleotides.[176] As DNA sequences increase in length and complexity, however, their eligibility for copyright protection would grow in proportion to their potential to be expressed in multiple ways.[177] Furthermore, DNA sequences having little or no functionality and abundant expression, such as the "Shotgun Wedding" puzzle in the 2005 MIT Mystery Hunt, would be readily eligible for copyright protection. Synthetic DNA would be relatively more likely than genomic DNA to qualify for copyright protection

---

L. REV. 1387, 1392–93 (2011); Thomas H. Murray, *What Synthetic Genomes Mean for our Future: Technology, Ethics, and Law, Interests and Identities*, 45 VAL. U. L. REV. 1315, 1321, 1338 (2011); Eleonore Pauwels, *Who Let the Humanists into the Lab?*, 45 VAL. U. L. REV. 1447, 1454–56 (2011).

176 *See supra* Part I (explaining the potential of DNA copyright protection).

177 Burk, *supra* note 7, at 501. Dan Burk uses the example of relatively unconstrained "enhancer control sequence," observing that "[i]n the case of enhancers, the same function may be achieved through many arrangements, and thus a particular arrangement may be copyrightable." *Id.* In a footnote to this sentence, Burk adds:

> A closer question may be presented where elements are not absolutely constrained by functional considerations, but simply arranged for the sake of efficiency. If only one or a few arrangements are most efficient, merger may again prevent their copyrightablility. Naturally, these would be the arrangements innovators would most want to protect.

Id.

because of its origin within a human design milieu rich in opportunities for expressive choices.[178]

F. Duration

A copyright term lasts substantially longer than a patent term. A valid patent only lasts from the date it was issued until 20 years from the United States filing date of its corresponding patent application.[179] For most works of authorship created on or after January 1, 1978, the copyright term for an individual author may last throughout the life of that author plus 70 additional years,[180] while the term for anonymous and pseudonymous works and works for hire last for the earlier of 75 years from publication or 100 years from creation.[181] The longer term of a DNA copyright would increase the value of this right to its owner. However, any deadweight loss caused by this monopoly right would create long-term costs to society. Several features of copyright law would act to lessen this burden, such as independent creation, fair use, and the opportunities DNA copyright would create for open source biology.

G. Independent Creation

Unlike the case in patent law, copyright law frees from liability independently created works, which are identical to copyrighted works. Copyright law simply requires independent creation.[182] As long as a work authored second in time is original and not copied, it does not infringe an identical work created first in time. In fact, both works would qualify for copyright protection. By contrast, even independently created inventions can infringe a patent under the prevailing regime of strict liability.

H. Fair Use

Unlike patent law, which applies a strict liability standard to instances of infringement offering few and insubstantial exceptions,

---

178 Note that DNA sequences generated by computer software and lacking expressive choices would have relatively less eligibility for copyright protection.

179 35 U.S.C. § 154(a)(2) (2006). There are minor variations to this term for patents that reference priority documents under 35 U.S.C. § 120, § 121, and § 365(c). In addition, the patent term may be extended under 35 U.S.C. § 154(b), § 155, § 155A, and § 156.

180 17 U.S.C. § 302(a). For joint authors, the term lasts 70 years beyond the death of the last author to die. 17 U.S.C. § 302(b).

181 17 U.S.C. § 302(c).

182 17 U.S.C. § 101 (2006).

copyright law includes a provision that explicitly allows several significant uses of copyrighted works without resulting in liability for infringement. Section 107 of the Copyright Act, entitled "Limitations on exclusive rights: Fair use," describes this safe harbor from copyright infringement.[183] It states:

> the fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by [§ 106 and § 106A], for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright.[184]

Several of these enumerated instances of fair use are highly relevant to DNA sequences. Section 107 would appear to contemplate the copying and use of copyrighted DNA sequences for educational purposes. For example, biology professors and their students would seem permitted to make "multiple copies" of copyrighted DNA sequences for use in their studies of genetics in their classrooms and teaching laboratories. In addition, scholars and researchers of genetics would appear to be able to make and use copies of DNA sequences without triggering liability for copyright infringement. For example, a geneticist could seemingly copy and use copyrighted genes for her research, while, by contrast, such activities would trigger strict liability under patent law if the genes in question were claimed in a patent.

In addition to enumerated examples of copying that would not constitute copyright infringement, § 107 requires a mandatory analysis of four factors to determine "whether the use made of a work in any particular case is a fair use."[185] These four factors are:

> (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit education purposes;
> (2) the nature of the copyrighted work;
> (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and

183 17 U.S.C. § 107 (2006).
184 *Id.*
185 *Id.*

(4) the effect of the use upon the potential market for or value of the copyrighted work.[186]

The relatively stronger case for fair use tends to occur when the copying of the work is only partial and the context is noncommercial. Examples might include when copyrighted DNA sequences are used transformatively in scholarship or the DNA sequences are copied during academic laboratory research. Alternatively, the relatively weaker case for fair use tends to occur when the copying of the work is complete and the context is commercial. Examples might include the wholesale representation of complete nucleotide sequences of copyrighted DNA molecules on a commercial website or the industrial replication of copyrighted DNA sequences for sale at a profit.

Another exception to copyright infringement can be found in § 108, which allows libraries and archives to make and lend a copy of a DNA sequence as long as it is done in accordance with the other provisions of the section.[187]

Patent law allows very few instances of copying to escape infringement liability. Even the use of patented inventions in the educational or research environments of universities can trigger patent infringement.[188] This strict liability regime may chill even noncommercial activities, such as academic research, that would seem to pose minimal economic threat to owners of patent rights. The fair use defense in copyright creates a significant safe harbor within which socially valuable activities, such as academic research, may survive and perhaps, even thrive. Many critics and scientists who consider current patent law too unforgiving to genetic research would welcome a copyright regime with a robust fair use exception to infringement.

I. Open Source Biology

Open source software has generated many valuable innovations. These include Linux operating systems, Apache server software, and Ruby on Rails database software. Much open source software is created under the rubric of an open source license that relies on copyright law to enforce its provisions. Though often proposed, open source biology has thus far failed to make much of an impact on the field of biology. In part, the failure of open source biology can be blamed on the difficulty of adapting the patent system to an open source license. Unless all patent rights are covered by an open source license—a virtual impossibility—

186 *Id.*
187 17 U.S.C. § 108 (2006).
188 Madey v. Duke, 307 F.3d 1351 (Fed. Cir. 2002).

open source biology developers will fear the existence of patent liability for their activities. However, were DNA sequences to be protected by copyright rather than patent law, open source licenses could potentially offer the same advantages to open source biology as they do to open source software. If synthetic biology were indeed to make it technically feasible to develop programmable genes and organisms, open source biology licenses undergirded by copyright for DNA sequences could help the field achieve its potential.

VI. CONCLUSION

Copyright law has traditionally afforded protection to works of authorship such as books, magazines, photographs, paintings, music, and sculpture. The Copyright Act has proved admirably flexible at accommodating novel categories of authorship, specifically contemplating future developments by covering "original works of authorship fixed in any tangible medium of expression, *now known or later developed*."[189] This has led to explicit copyright protection for nontraditional works of architecture and computer software. Sequences of DNA should also be acknowledged as eligible for copyright protection.

Unaltered genomic DNA sequences would seem poor candidates for copyright protection. The case is stronger for copyright protection of recombinant DNA sequences. Strongest is the case for the copyright eligibility of synthetic DNA sequences designed nucleotide by nucleotide and chemically constructed *de novo*. Whereas DNA copyright has previously remained a largely hypothetical prospect, advances in synthetic biology may now force recognition of copyright protection as an alternative (or complement) to patent protection.

A DNA copyright regime would differ substantially from the current DNA patent regime. Notably, acquiring copyright protection for DNA would be less expensive and much more rapid than pursuing patent protection. As patent law recognizes few and weak exceptions to infringement, copyright law offers a robust fair use exception for copying done in contexts such as scholarship and research. Furthermore, copyright protection would be limited in the case of DNA molecules whose structures are dictated by functional constraints, thus providing the public greater and salutary access to useful genes.

Copyright protection for DNA lies pregnant within current copyright law. What is required is an effort to make use of the existing protection. A DNA copyright regime would not only allow a more

---

[189] 17 U.S.C. § 102 (2006) (emphasis added).

robust set of safe harbors for use of particular DNA sequences, especially in genetic research, it would also facilitate the possibility of an open source biology movement. Finally, just as the prospects of patent protection for at least some forms of DNA have become uncertain, copyright protection could fill any resulting gap by affording a reasonable level of intellectual property protection, while simultaneously allowing society to enjoy some of the benefits of genetic knowledge more freely than the current patent protection is able to afford.