

Are Engineered Genetic Sequences Copyrightable?: The U.S. Copyright Office Addresses a Matter of First Impression

By CHRISTOPHER M. HOLMAN, CLAES GUSTAFSSON, and ANDREW W. TORRANCE

IN THE EARLY 1980s, at a point in time when modern biotechnology was just coming into its own, an attendee at a continuing legal education (CLE) program on copyright law asked a law professor named Irving Kayton to explain why genetically engineered DNA sequences could not be copyrighted. Although initially “shocked and perplexed” by the suggestion, he decided to pursue the question and ultimately concluded that “[u]nder certain circumstances, from a practical as well as legal viewpoint, copyright protection may be the only or the most effective way an ‘author’ can protect a valuable genetic ‘work.’”¹ In going through the exercise, Kayton confessed that “every intellectual and emotional prejudice, both sophisticated and primitive, to which [a person] is subject opposed coming to the conclusion finally reached. Copyright for engineered DNA sequences seemed ludicrous.” Kayton’s conclusion that engineered genetic sequences are copyrightable under the Copyright Statute was based largely on the analogy between engineered genetic code and computer code, and the fact that even highly functional computer programs can be copyrighted.²

In the years following Kayton’s article, a number of legal scholars, including two authors of the present article, have observed that the argument in favor of extending copyright to engineered DNA sequences has only gotten stronger as the historically distinct disciplines of synthetic biology and soft-

ware engineering increasingly converge.³ Stanford’s Professor Drew Endy recently opined that “given the history in software, there is going to be for the foreseeable future an ever-renewing enthusiasm for exploring the idea of copyright” for synthetic biology.⁴ He noted that “literally every student I see ... who connects with property rights immediately presumes that you should be treating this stuff like code, and they are familiar with using copyright in that context.”⁵

In spite of the compelling logic that would support extending copyright to engineered DNA sequences, copyright protection for genetic code has not been legally recognized in the U.S., or as far as we know anywhere. The Copyright Act is silent on the point, the courts do not appear to have ever addressed the question, and the Copyright Office

¹Irving Kayton, *Copyright in Living Genetically Engineered Works*, 50 GEO. WASH. L. REV. 191, 192 (1982).

²*Id.*

³See, e.g., 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 (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); Christopher M. Holman, *Copyright for Engineered DNA: An Idea Whose Time Has Come*, 113 W. VA. L. REV. 699–738 (2011); Andrew W. Torrance, *Synthesizing Law for Synthetic Biology*, 11(2) MINN. J. L. SCI. & TECH. 629–665 (2010); Andrew W. Torrance, *DNA Copyright*, 46(1) VALPARAISO U. L. REV. 1–41 (2011).

⁴Drew Endy, Professor, Stanford Univ., Keynote Address at the Stanford Law School Conference on Intellectual Property Law and the Biosciences (Apr. 27, 2012), http://www.youtube.com/watch?v=Qku3OQ5O_U4, at 50:30–52:00 minutes (last visited Jan. 11, 2016).

⁵*Id.*

Christopher M. Holman is the Executive Editor of *Biotechnology Law Report* and a Professor at the University of Missouri–Kansas City School of Law. Claes Gustafsson is the Chief Commercial Officer and a Co-Founder of DNA 2.0 in Menlo Park, California. Andrew W. Torrance is a Professor at the University of Kansas School of Law in Lawrence, Kansas.

has taken the position that an engineered genetic sequence is not copyrightable subject matter.⁶ In an attempt to advance the conversation, we submitted an engineered DNA sequence to the Copyright Office for registration, and then appealed the Office's decision refusing to register engineered genetic sequences. This article reports the outcome of our experiment, and provides as supplementary material the actual letter we submitted to the Copyright Office appealing its initial decision not to register genetic sequences (the "Appeal"),⁷ along with the Copyright Office's letter denying our appeal (the "Denial"),⁸ which provides a detailed explanation of the Office's position regarding the copyrightability of engineered DNA. The bulk of the article is devoted to refuting the legal and policy justifications set forth by the Office in its Denial.

THE RATIONALE BEHIND OUR DECISION TO SEEK REGISTRATION OF AN ENGINEERED GENETIC SEQUENCE

A number of possible paths exist by which U.S. law might come to recognize the copyrightability of engineered genetic code. Congress could amend the Copyright Statute to explicitly identify engineered genetic code as a category of copyrightable subject matter, but as a practical matter this seems unlikely, at least in the near term. Alternatively, the courts could recognize the copyrightability of engineered genetic code, without requiring any explicit congressional action.

There are essentially two scenarios whereby the courts could declare engineered genetic sequences copyrightable. One would be in the context of an infringement litigation. Significantly, unlike a patent, there are no formalities required for copyright protection, *i.e.*, a work does not have to be registered with the Copyright Office in order to be afforded copyright protection.⁹ Infringement lawsuits were the primary mechanism whereby copyright protection for software was recognized, but as of yet no one appears to have filed a lawsuit claiming infringement of copyright on an engineered genetic sequence (although there are reports that law firms and biotechnology companies have begun discussing the possibility).

The other scenario would be a judicial appeal of a decision by the Copyright Office to deny registration of an engineered DNA sequence. This was the route through which the copyrightability of a video game display, for example, was established. In that case, the Copyright Office had repeatedly

rejected Sega's attempts to register the video game Breakout as an audiovisual work, based on the Office's conclusion that "the display screens both individually and as a whole simply lack[] sufficient creativity to make them registerable as audiovisual works."¹⁰ On appeal, the D.C. Circuit found that the Copyright Office had erred by applying an overly stringent standard of creativity, emphasizing that "[i]t is not the Register's task to shape the protection threshold or ratchet it up beyond the 'minimal creative spark required by the Copyright Statute and the Constitution.'"¹¹

The recognition of copyright protection for software only occurred after many years of serious debate and discussion, and it seems likely that it will be the same with engineered DNA sequences. In an attempt to move the ball forward, the two law professor authors of this article (Holman and Torrance) teamed up with DNA 2.0, a leading gene synthesis and design company, to attempt the registration of a synthetic DNA sequence created by the company. The sequence, dubbed "Prancer," encodes a non-naturally occurring fluorescent protein.

On July 3, 2012, DNA 2.0 filed a request for registration of the Prancer sequence, and the following month the Copyright Office responded with a generic form letter refusing registration of the DNA sequence. The letter provided no specific explanation for the refusal, stating only that the "material submitted does not contain the minimum amount of authorship required for registration."

In pursuit of a more full explanation, we submitted a request for reconsideration (the "Appeal") on

⁶COMPENDIUM OF U.S. COPYRIGHT OFFICE PRACTICES § 313.3(A) (3rd ed. Dec. 22, 2014).

⁷DNA 2.0, *Request for Reconsideration of Denial of Copyright Registration of Prancer DNA Sequence* (Nov. 26, 2012) (referred to herein as the "Appeal") [provided as Supplementary Document 1].

⁸Robert J. Kasunic, Associate Register of Copyrights and Director of Copyright Policy and Practices, *Affirmance of Refusal for Registration* (Feb. 11, 2014) (referred to herein as the "Denial") [provided as Supplementary Document 2].

⁹*Paul Morelli Design, Inc. v. Tiffany and Co.*, 200 F. Supp. 2d 482, 484 (E.D. Pa. 2002) (In order to sue for copyright infringement, a party must either (1) have obtained a registration of the copyright from the Register of Copyrights in the Library of Congress or (2) have applied for a registration and had the registration refused by the Register. 17 U.S.C. § 411(a).) *Atari Games Corp. v. Oman*, 888 F.2d 878, 880–81 (D.C. Cir. 1989).

¹⁰*Atari Games Corp. v. Oman*, 979 F.2d 242, 243 (D.C. Cir. 1992).

¹¹*Id.*

November 26, 2012.¹² In our request, we argued that human-designed DNA sequences such as Prancer fall comfortably within the category of “literary work” explicitly specified as copyrightable in the statute, for substantially the same reasons that computer programs are currently treated as literary works eligible for copyright protection. Our Appeal explained that as an original work of authorship “fixed in [a] tangible medium of expression,” the Prancer DNA sequence appears to satisfy the various statutory requirements of copyright, particularly given the copyright statute’s expansive and flexible definition of copyrightable subject matter. We also emphasized the potential policy benefits that would accompany a recognition of copyright protection for engineered genetic code.

In February 2014, we received a letter from the Copyright Office responding to our request for reconsideration (the “Denial”).¹³ The Denial, signed by the Associate Register of Copyrights and Director of Copyright Policy and Practices, began by apologizing for the delay in responding to our request, explaining that “this request was an issue of first impression for the U.S. Copyright Office and as such, was given significant consideration prior to rendering a decision.” The letter goes on to state that “after carefully reconsidering the registration materials and the arguments contained in your request reconsideration, the Office affirms the refusal of registration.”

The Denial sets forth both policy and legal rationales purporting to support the Office’s decision to refuse registration. We could have petitioned the Copyright Office a second time to reconsider its refusal to register, and then proceeded to challenge the decision in the courts, as Sega did with respect to videogame displays discussed above. However, after reading the Denial we concluded that further appeal to the Copyright Office would almost certainly be futile, and an appeal to the courts would require the expenditure of more time and money than DNA 2.0 was at that time prepared to spend. However, we decided to publish this article in order to bring more attention to our efforts and to publicly address, and we believe refute, the policy and legal rationales put forward by the Copyright Office for denying the copyrightability of engineered DNA sequences.

REFUTING THE COPYRIGHT OFFICE’S REASONS FOR DENIAL

This section of the article addresses and refutes the primary justifications asserted by the Copyright Office in support of its decision to deny registration to engineered genetic sequences.

The Office’s assertion that engineered genetic sequences do not fall within an explicitly enumerated category

The Copyright Office’s primary argument against registering engineered genetic sequences appears premised on its assertion that genetic sequences do not fall within any of the eight categories of copyrightable subject matter explicitly enumerated in Section 102(a) of the Copyright Statute, and that neither the Office nor the courts have the authority to declare copyrightable subject matter that does not fall within one of these eight categories. But this interpretation of the law is entirely inconsistent with both the language and the legislative history of the statute. While it is true that Section 102(a) enumerates eight categories of copyrightable subject matter, including, for example, literary works, musical works, and dramatic works, it explicitly states that “works of [copyrightable] authorship *include* the [enumerated] categories.” In other words, the eight enumerated categories are illustrative, not limiting, and a work does not necessarily have to fall into one of these categories to be afforded copyright protection. This has long been the conventional understanding of copyright attorneys and scholars. For example, the leading copyright treatise, *Nimmer on Copyright*, explains that it is “clear that ‘works of authorship’ are not necessarily limited to the eight broad categories of works listed under Section 102(a).”¹⁴ *Nimmer* points to the legislative history of the Copyright Act, which explicitly states that these categories are “‘illustrative and not limitative,’ and ... do not necessarily exhaust the scope of ‘original works of authorship’ that the bill is intended to protect.”¹⁵

Although it seems clearly incorrect for the Office to assert that it is precluded by statute from registering subject matter that does not fall within one of the enumerated categories, one can understand why the Office, as a matter of policy, might adopt this position. In 1980, the Supreme Court faced a similar question in *Diamond v. Chakrabarty*, when it was called upon to decide whether genetically engineered living organisms could be patented.¹⁶ Several dissenting Justices argued that the Court should not take the lead in expanding the scope of patentable subject matter to encompass new inventions

¹²Appeal, *supra* note 7.

¹³Denial, *supra* note 8.

¹⁴1–2 NIMMER ON COPYRIGHT § 2.03.

¹⁵*Id.*

¹⁶*Diamond v. Chakrabarty*, 447 U.S. 303 (1980).

arising out of biotechnology, but that instead the decision to expand patent law to biotechnology should be left to Congress. The majority rejected this approach, however, and held that Congress intended patent law to be dynamic and capable of expanding to accommodate important new developments in technology such as biotech. In retrospect, many people point to *Chakrabarty* as a pivotal event that helped encourage investment in biotechnology by reassuring investors that patent protection would be available for resulting innovations. Significantly, Congress has never explicitly authorized the extension of patents to new technologies like biotechnology, but instead has permitted the courts to oversee the expansion of patentable subject matter to accommodate new technologies.

Historically, copyright law has also expanded to encompass new technologies without explicit congressional action. Software is a good example. Today the availability of copyright protection for computer programs is well established, but the Copyright Statute has never been explicitly amended to provide for the copyrightability of software. Instead, it has largely been the courts that have overseen the expansion of copyright to encompass even highly functional software products written in object code and fixed in media that can only be interpreted by a computer.

The legislative history of the Copyright Statute indicates that Congress intended to allow for its expansion to accommodate new technologies. The law has a long and well-established tradition of using analogy as a primary mechanism for adapting to the development of new technologies that, in view of the overriding policy objectives underlying copyright law, warrant copyright protection. For example, Professor Nimmer has observed that “[a]s to new forms of creative expression that may emerge in the future as a result of scientific discoveries or technological developments ... [i]f such a new form is sufficiently analogous to the kinds of works that are expressly protected in the eight categories, it will be regarded as falling within ‘the present congressional intent,’ even though the similarity is only by analogy.”¹⁷ Similarly, in his treatise, Professor Goldstein correctly observes that “the question will sometimes arise whether a new form of authorship, not expressly mentioned in the Act, is entitled to protection. The most practical and principled approach to this question is to reason by analogy to works expressly listed in section 102. ... new forms of works should be protected if they are similar to those listed and not protected if they are dissimilar.”¹⁸

As one of us has pointed out in a previous article, the analogy between software and engineered ge-

netic code is striking and becoming ever more so as the two fields evolve and converge.¹⁹ A 2010 article, co-authored by a group of biologists, computer scientists, and computational biologists, makes this point nicely:

Given that counting genes in the genome is such a large-scale computational endeavor and that genes fundamentally deal with information processing, the lexicon of computer science naturally has been increasingly applied to describing them. In particular, people in the computational biology community have used the description of a formal language to describe the structure of genes in very much the same way that grammars are used to describe computer programs—with a precise syntax of upstream regulation, exons, and introns. Moreover, one metaphor that is increasingly popular for describing genes is to think of them in terms of subroutines in a huge operating system (OS). That is, insofar as the nucleotides of the genome are put together into a code that is executed through the process of transcription and translation, the genome can be thought of as an operating system for a living being. Genes are then individual subroutines in this overall system that are repetitively called in the process of transcription.²⁰

Nonetheless, the Denial rejected our arguments with respect to the close analogy between engineered genetic code and computer code. The Office’s primary justification for its divergent treatment of genetic and computer code seems to be based on Congress’s decision in 1980 to amend the Copyright Statute to include a definition for the term “computer program.” The Copyright Office appears to believe that in making this amendment, Congress intended to expand the scope of copyrightable subject matter to include computer programs. Thus, according to the Denial, computer programs are to be treated differently from other new forms of expression made possible by advances

¹⁷1–2 NIMMER ON COPYRIGHT § 2.03.

¹⁸GOLDSTEIN ON COPYRIGHT § 2.6 (2:86) (citing House Report on the 1976 Act at 51).

¹⁹Christopher M. Holman, *Developments in Synthetic Biology Are Altering the IP Imperatives of Biotechnology*, 17 VANDERBILT J. ENT. & TECH. L. 385 (2015).

²⁰Mark B. Gerstein *et al.*, *What Is a Gene, Post-ENCODE? History and Updated Definition*, 17 GENOME RES. 669–681 (2007).

in technology, no matter how high the degree of analogy between computer programs and other new forms of expression, like engineered genetic code. In our view, the Office's position reflects a misunderstanding of the significance of the 1980 amendment.

The context of the amendment is significant in this regard. In 1980, Congress made two amendments to the Copyright Statute relating to computer programs. First, it completely rewrote Section 117, renaming it as "Limitations on Exclusive Rights: Computer Programs." This new Section 117 created certain limitations on the enforceability of copyright on computer programs. Second, Congress introduced a definition of the term "computer programs" into Section 101 of the statute, presumably in an attempt to reduce ambiguity with respect to the proper interpretation of new Section 117. Significantly, Congress has never amended the Copyright Statute to explicitly extend copyright protection to computer programs. The most plausible explanation for this is that Congress believed that computer programs were already copyrightable under the Copyright Statute. Indeed, this was the conclusion of the CONTU Report, which was written by a panel of experts established by Congress in the 1970s to study and make recommendations as to how the copyright law should respond to the increasing commercial significance of computer programs.²¹

If Congress had not believed that the copyright statute already encompassed computer programs, it could, and presumably would, have amended the statute to explicitly cover computer programs. For example, as originally enacted in 1976, Section 102(b) of the statute only enumerated seven illustrative categories of copyrightable subject matter.²² When Congress decided in 1990 to include "architectural works" within the realm of copyrightable subject matter, it explicitly amended Section 102(b) by adding "architectural works" as an eighth enumerated category.²³ Thus, while the Copyright Office treats the 1980 amendments as evidence that Congress wished to limit the expansion of copyright to computer programs, a better interpretation is that the amendments evidence the fact that Congress viewed 102(b) as providing a broad and expansive, non-limiting definition of copyrightable subject matter that, as a general matter, encompasses new forms of expression made possible by advances in technology. The Appeal provides no explanation as to why computer code should be treated any differently than genetic code. In fact, most of the important justifications for extending copyright to computer programs cited in the CONTU Report apply equally to engineered genetic sequences.²⁴

It bears noting that the Office's current stance regarding the registration of engineered genetic sequences is entirely inconsistent with the approach it took with respect to software in the 1960s. People began attempting to register computer programs in the early 1960s, and the Copyright Office initially expressed "profound doubts" as to whether computer programs qualified as copyrightable subject matter.²⁵ Nonetheless, in 1964 the Copyright Office began to permit registration of computer programs under its "rule of doubt[], leaving the ultimate question of copyrightability to the courts."²⁶

Furthermore, the Denial provides no convincing rationale for the Office's decision to treat computer code, but not genetic code, as a form of literary work. Admittedly, the characterization of computer programs as "literary works" is quite a stretch,²⁷ but the stretch occurred decades ago, when the courts and Copyright Office began treating highly functional computer programs, written in object code only interpretable by a computer, as literary works. Given the growing analogy between engineered genetic code and computer code, there is little if any additional stretch in going from computer code to engineered genetic sequences.

The Denial argues that the use of a sequence of letters to represent the Prancer DNA sequence does not constitute "literary authorship, but rather a form of notating the biological sequence," because "[e]very gene sequence is represented by some

²¹National Commission on New Technological Uses of Copyrighted Works, Final Report, at 15–16 (1979) (often referred to as the "CONTU Report").

²²Pub. L. No. 94-553 (S. 22), 90 Stat. 2541 (Oct. 19, 1976).

²³Judicial Improvements Act of 1990, Pub. L. No. 101-650, 104 Stat. 5089 (Dec. 1, 1990).

²⁴Christopher M. Holman, *Copyright for Engineered DNA: An Idea Whose Time Has Come*, 113 W. VA. L. REV. 699–738 (2011).

²⁵Pamela Samuelson, *CONTU Revisited: The Case Against Copyright Protection for Computer Programs in Machine-Readable Form*, 1984 DUKE L.J. 663, 692–93 (1984) (citing George Cary, *Copyright Registration and Computer Programs*, 11 BULL. COPYRIGHT SOC'Y 362, 363 (1964)).

²⁶*Id.* See also William F. Patry, *Copyright and Computer Programs: A Failed Experiment and a Solution to a Dilemma*, 46 N.Y.L. SCH. L. REV. 201 (2003).

²⁷It is interesting to note that a district court judge recently issued an order indicating that, in her opinion, computer programs are not literary works. *Adobe Sys. Inc. v. A&S Electronics, Inc.*, No. 15-cv-02288-SBA, slip op. at 6–7 (N.D. Cal. Aug. 19, 2015) ("[17 U.S.C. § 106(4)] expressly applies only to 'literary, musical, dramatic, and choreographic works, pantomimes, and motion pictures and other audiovisual works,' not to computer software.").

sequence of these four letters to identify the sequence of the chemical compounds that these letters represent.” But once again, the Office fails to explain how this is any different than the case of software. It is well established that even a computer program written in object code, *i.e.*, as a string of zeros and ones interpretable only by a computer, is copyrightable subject matter. But this string of zeroes and ones is nothing more than “notation” (to use the language of the Denial) that identifies, in the case of an optical storage device such as a CD-ROM or DVD, a sequence of deformities on the surface of a circular disc, or, in the case of a magnetic storage device such as a floppy disk or hard drive, a pattern of magnetization on a magnetically coated surface. The fact that computer code and genetic code “notate” the physical attributes of a medium used to convey information to a machine or device does not render the underlying code uncopyrightable *per se*.

The Office’s inability to perform a search for similar sequences

The Denial asserts that the Copyright Office lacks the ability to perform a search of naturally occurring DNA sequences in order to determine whether a sequence submitted for registration is derived from a naturally occurring sequence. This is essentially a policy, as opposed to legal, basis for refusing registration. What is not clear is why the Office would feel it is necessary to perform such a search. As a general matter, the Copyright Office does not perform searches to determine if subject matter that has been submitted for registration is identical or similar to pre-existing works.²⁸ Moreover, the Office presumably lacked the ability to perform searches with respect to computer programs it registered in the 1960s under the rule of doubt. The Denial provides no explanation as to why there is a unique need to perform a search with respect to an engineered genetic sequence submitted for registration.

Furthermore, if the Copyright Office deems it necessary or desirable to perform searches on engineered genetic sequences submitted for registration, then it would seem incumbent upon the Office to attain that capability, rather than denying registration to copyright-eligible subject matter. The Patent and Trademark Office (PTO) has not always been equipped to search for prior art in the realm of DNA sequences (or software for that matter), but as inventors began applying for patents on these technologies, the PTO responded by developing the capabilities.²⁹

A group of copyright experts recently recommended that the Copyright Office consider outsourcing some registration responsibilities to third parties as a means of addressing certain limitations in the capabilities of the Office.³⁰ Registration of engineered genetic sequences could be an area where such outsourcing would be particularly useful, particularly if the Office concludes that it is important to perform searches prior to registration. For example, the Copyright Office could delegate DNA sequence searches to the PTO, or some other agency with more expertise in searching and examining genetic sequences, or even to a private third party.³¹

Overlapping copyright and patent protection

The Denial goes on to conclude that the “inability of the Office to independently discern new creative authorship [*i.e.*, to perform a prior art search] suggests that a claim in a DNA sequence may be far better suited for the realm of patent, where a heightened standard of novelty, nonobviousness, and an examination of prior art would be considered, rather than the originality standard of copyright.” It goes on to assert that the fact that engineered DNA sequences are patent eligible “provides reason to question whether synthetic or cDNA

²⁸COMPENDIUM OF U.S. COPYRIGHT OFFICE PRACTICES § 602.4(C), *No Searches or Comparison of Works* (3rd ed. Dec. 22, 2014) (“When examining a claim to copyright, the U.S. Copyright Office generally does not compare deposit copy(ies) to determine whether the work for which registration is sought is substantially similar to another work. Likewise, the Office generally does not conduct searches to determine whether the work has been previously registered.”).

²⁹MANUAL OF PATENT EXAMINING PROCEDURE § 2420, *The Requirements for Patent Applications Containing Nucleotide Sequence and/or Amino Acid Sequence Disclosures—the Sequence Rules* [R-08.2012] (8th ed. 2001) (describing difficulties in searching and examining nucleotide sequences prior to the implementation of rules relating to nucleotide and amino acid sequence disclosures).

³⁰Pamela Samuelson, *The Copyright Principles Project: Directions for Reform*, 25 BERKELEY TECH. L.J. 1175 (2010), UC Berkeley Public Law Research Paper No. 1851857, available at SSRN: <http://ssrn.com/abstract=1851857> (“The Copyright Office should transition away from being the sole registry for copyrighted works and toward certifying the operation of registries operated by third parties, both public and private.”).

³¹Chris Holman, *Unlocking DNA Copyright Protection with Sequence Searches (Part 1)*, GQ LIFESCIENCES BLOG (May 2, 2016), available at <https://www.gqlifesciences.com/unlocking-dna-copyright-protection-with-sequence-searches-part-1/> (last visited May 14, 2016).

sequences are proper subject matter for copyright,” pointing to a “potential overlap between copyright and patent protection.” But there is absolutely no reason why both forms of protection cannot overlap with respect to the same subject matter. The Federal Circuit recently rejected any suggestion that the availability of patent protection for software should render it ineligible for copyright protection, emphasizing that the Supreme Court has made clear that “[n]either the Copyright Statute nor any other says that because a thing is patentable it may not be copyrighted.”³²

The Denial also makes much of the fact that “Congress has chosen to provide delineated patent protection for certain biological processes, such as plant patents for newly invented strains of asexually reproducing plants.” But the suggestion that the availability of plant patents somehow negates the availability of alternate forms of intellectual property protection for biotechnology was effectively refuted by the Supreme Court in *J.E.M. Ag Supply, Inc. v. Pioneer Hi-Bred Int’l*.³³

The Office’s assertion that engineered sequences are uncopyrightable under 17 U.S.C. Section 102(b)

The Office goes on to assert that genetic code is uncopyrightable under Section 102(b) of the copyright statute, which prohibits the extension of copyright to “any idea, procedure, process, system, method of operation, concept, principle, or discovery.” According to the Denial, “The Prancer DNA Sequence is genetic formula for a biological system. It does not describe, explain, or illustrate anything except the genetic markers that comprise this biological organism. Therefore, there is no copyrightable expression, but rather the claim simply records the formula for this biological system or process.”

At this point the Office seems to have entirely lost sight of the nature of engineered DNA. The reference to “genetic markers” and “this biological organism” make no sense—an engineered genetic sequence does not constitute a genetic marker, nor is it a biological organism. The characterization of the sequence as a “genetic formula for a biological system” is likewise difficult for a biologist to interpret, and appears to reflect a lack of understanding.

In any event, the analogous argument that copyright for software is precluded by Section 102(b) has been raised from time to time, and repeatedly rejected by the courts. For example, in *Oracle v. Google*, the court held that “components of the program that can be characterized as a ‘method of operation’ may nevertheless be copyrightable. The

Oracle court also pointed out that “classifying a work as a ‘system’ does not preclude copyright for the particular expression of that system.”

The functionality and alleged artistic deficiencies of engineered sequences

The letter acknowledges that “[i]n the case of a synthetic gene, the specific sequence of nucleotides is the result of some person’s choices,” but goes on to conclude that these choices are not sufficient to render the resulting sequence copyrightable because “those choices are not made for the purpose of artistic expression. . . . Properly understood, the nucleotide sequence of a synthetic gene, inasmuch as it could be conceived as a form of expression at all, is a form of expression dictated solely by functional considerations.”

But it is incorrect to suggest that copyright protection is limited to “artistic” works, and denied to functional works. It is well established that even highly functional software can be copyrighted, regardless of whether it incorporates any artistic element that could be discerned by the Copyright Office. Indeed, in *Oracle*, the court recognized that “computer programs are by definition functional,” and that Congress and judicial precedent firmly establish that computer programs are copyrightable, “despite their utilitarian or functional purpose.”³⁴ Of course, one might argue that the author of even the most functional computer program might incorporate elements that would be perceived as aesthetically pleasing by other software designers, but the same could be said for the design of a synthetic genetic sequence. As an aside, there are examples of non-functional synthetic DNA sequences that would easily satisfy even the “artistic” creativity threshold proposed by the Office, including literary quotations introduced into a synthetic *M. genitalium* genome by Craig Venter’s group and a version of George Church’s book, *Regenesis—How*

³²*Oracle Am., Inc. v. Google Inc.*, 750 F.3d 1339, 1380–81 (Fed. Cir. 2014) (citing *Mazer v. Stein*, 347 U.S. 201, 217 (1954)).

³³*J.E.M. Ag Supply, Inc. v. Pioneer Hi-Bred Int’l, Inc.*, 534 U.S. 124, 132 (2001) (The 1930 PPA [Plant Patent Act] conferred patent protection to asexually reproduced plants. Significantly, nothing within either the original 1930 text of the statute or its recodified version in 1952 indicates that the PPA’s protection for asexually reproduced plants was intended to be exclusive.).

³⁴750 F.3d at 1367–1368 (“Section 102(b) does not bar the packages from copyright protection just because they also perform functions”).

Synthetic Biology Will Reinvent Nature and Ourselves, encoded in DNA.³⁵

*A bias against biological creativity
and works based on nature*

The Denial characterizes engineered DNA as the “result of . . . biological techniques,” and asserts that “[b]iological creations do not fit within any of the existing categories of authorship.” It goes on to state that it is unclear whether a DNA sequence can ever be “entirely human-created,” noting that “[s]ince the operation of DNA is dictated by laws of biology, at least some aspects of DNA sequences are controlled by those laws.” Without further explanation from the Office, it is hard to know exactly what to make of such assertions. To say that the operation of DNA is dictated by laws of biology seems comparable to saying that the operation of software is dictated by the laws of physics, which is of course true, but does not preclude copyrightability.

The Copyright Office seems to be expressing a bias against the idea of copyright extending into the realm of biology. Professor Kayton himself confessed to the same sort of initial prejudice when he was first confronted with the idea of copyright for engineered genetic sequences. His prejudice was only overcome after he grappled with the issue for some time and was forced to acknowledge that logic dictated that copyright protection for computer code imputes copyright protection for genetic code.

Similarly, in response to our argument that DNA is analogous to a computer program because it is a set of instructions directed towards a biological machine, the Office asserts that “[w]hile an organism may be analogized to a machine, it is clearly not one, and as a result falls outside of the category enumerated in the statute”; *i.e.*, it is not a literary work. The Copyright Statute specifies that copyright protection is available for “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.”³⁶ So it appears that the Office is arguing that the determinative difference between computer code and genetic code is that computer code is directed towards a computer, which is a “machine or device” as explicitly provided for in the statute, while genetic code is directed towards a biological organism or system which falls outside the bounds of “machine or device.”

Once again, this seems to be a clear case of discrimination against biological-based innovation.

The Copyright Statute explicitly defines the terms “machine or device” as broadly encompassing machines and devices “now known or later developed.”³⁷ The Merriam-Webster online dictionary defines a “device” as “an object, machine, or piece of equipment that has been made for some special purpose.” A “machine” is defined as “a piece of equipment with moving parts that does work when it is given power from electricity, gasoline, etc.” Both definitions, and particularly the definition of device, would appear to encompass a biological organism programed to perform some useful function, especially given that the Copyright Statute explicitly anticipates expansion of the definition of device and machine to keep up with ongoing advances in technology.

In fact, synthetic biologists often refer to engineered biological organisms as “machines,” and at times even as “computers.” For example, researchers at Stanford University recently reported using engineered DNA to create inside living cells a “biological device [that] behaves like a transistor, one of the tiny switches that are etched on to microchips in the billions to perform computer calculations.”³⁸ In an article published in the scientific journal *Science*, these researchers explained “how their biological transistors could be connected together inside living cells to perform computing jobs such as controlling how genes are expressed in an organism.”

The Denial goes on to state that “the extent to which any genetic sequence can truly be entirely human-created is unclear, particularly with respect to copyright creativity,” because “to the extent that any sequence involves the isolation of any naturally-occurring sequence, or a derivation thereof, it may not be deemed created as opposed to discovered.” In essence, the Office seems to be suggesting that the fact that engineered genetic sequences are, at

³⁵Press Release, J. Craig Venter Institute, *First Self-Replicating Synthetic Bacterial Cell* (May 20, 2010), available at <http://www.jcvi.org/cms/press/press-releases/full-text/article/first-self-replicating-synthetic-bacterial-cell-constructed-by-j-craig-venter-institute-researcher/home/> (last visited May 14, 2016); R. Alan Leo, *Writing the Book in DNA*, HARVARD MEDICAL SCHOOL (Aug. 16, 2012), available at <http://hms.harvard.edu/news/writing-book-dna-8-16-12> (last visited May 14, 2016).

³⁶17 U.S.C. § 102(a).

³⁷17 U.S.C. § 101.

³⁸Ian Sample, *Scientists Create Transistor-Like Biological Device*, THE GUARDIAN (Mar. 28, 2013), available at <http://www.theguardian.com/science/2013/mar/28/transistor-biological-device> (last visited May 14, 2016).

some level, based upon a modification or reconfiguration of genetic information ultimately derived from naturally occurring sequences somehow renders engineered genetic sequences uncopyrightable products of nature. But there is no prohibition against copyright protection for an expressive work based upon nature. The photography of Ansel Adams, capturing the natural beauty of Yosemite Valley, for example, is not precluded from the protection of copyright.

In *Proline Concrete Tools, Inc. v. Dennis*, a district court recently chastised the Office for making a similar mistake when it denied registration for “rock and stone sculptures ... used to make decorative concrete stamps.”³⁹ The decision to deny was based on the Office’s conclusion that the sculptures were “slavish copies of un-copyrightable objects and, as such, do not contain a sufficient amount of original authorship to support copyright claims.” The court overturned the Office’s decision, however, characterizing it as the result of “a clear misunderstanding on the part of the Register that seemed “to result from its confusion over how [the registrant] creates the sculptures and what exactly it seeks to copyright.”⁴⁰

Similarly, in the case of *Prancer*, the Office seems to misapprehend the nature of genetic engineering and has conflated the copyrightability of a naturally occurring genetic sequence, which no one is arguing should be eligible for copyright, with the copyrightability of an engineered genetic sequence. Like the stone sculptures in *Proline Concrete Tools*, the *Prancer* sequence is not a slavish copy of a naturally occurring genetic sequence, and registration of its sequence would in no way threaten the public’s access to any naturally occurring sequence. In large

part, the denial of registration appears to be based on the Office’s unwillingness to recognize the fundamental distinction between naturally occurring sequences and a sequence that diverges significantly from any naturally occurring sequence from which it may have been derived.

CONCLUSION

The Copyright Office’s position that engineered genetic sequences are not copyrightable seems to be based primarily on policy considerations, rather than an objective application of U.S. copyright law, and as such it seems unlikely that the Office will be persuaded solely by legal or technical arguments. Nonetheless, there are signs of a growing recognition amongst biotechnology stakeholders of the potential benefits of extending copyright protection to at least some engineered genetic sequences that meet a certain threshold of creativity and originality. The authors have been contacted by biotechnology companies and law firms who have expressed interest in our work, and it seems inevitable that at some point the question of copyright for engineered DNA will be before the courts, and perhaps ultimately addressed by Congress. We hope that by sharing our experience of attempting to register an engineered DNA sequence we have advanced the discussion, and that one day, the creators of original genetic code will be able to benefit from the same legal protection currently afforded to the creators of computer code.

• • •

³⁹*Proline Concrete Tools, Inc. v. Dennis*, No. 07CV2310-LAB AJB, 2012 WL 2886953, at *1 (S.D. Cal. July 13, 2012).

⁴⁰*Id.*

Supplementary Document 1: Request for Reconsideration of Denial of Copyright Registration of Prancer DNA Sequence



Via Express Messenger

November 26, 2012

US Copyright Office
Receipt Analysis and Control Division
PO Box 71380
Washington DC 20024-1380

Re: Request for Reconsideration of Denial of Copyright Registration of Prancer DNA Sequence

The United States Copyright Office has refused to register a DNA sequence referred to herein as the Prancer DNA Sequence, based on its determination that Prancer “does not contain the minimum amount of authorship required for registration.” Applicant seeks reconsideration based on the significant amount of authorship embodied in Prancer—an amount well exceeding the minimal standard for originality set forth by the courts. As the United States Supreme Court has stated, “the requisite level of creativity is extremely low; even a slight amount will suffice.”¹ Prancer easily satisfies this criterion.

Applicant recognizes that the Copyright Office in all likelihood has adopted a *de facto* position that DNA sequences are not copyrightable. Although the copyright eligibility of a human-engineered DNA sequence apparently is a legal issue of first impression with respect to the courts, human-designed DNA sequences such as Prancer fall comfortably within the category of “literary work” explicitly specified as copyrightable in the statute, for substantially the same reasons that computer programs currently are treated as literary works eligible for copyright protection.

It is important to emphasize that Prancer is a synthetic, human-designed DNA sequence. The Prancer DNA sequence is not derived from a natural source, and would not exist but for its human authorship. Just as in the case of a copyrightable computer program, the Prancer DNA sequence is an original work of authorship, is expressive, and has been fixed in a tangible medium of expression. A number of legal scholars have addressed the issue and, for a variety of reasons, have concluded that at least some human-designed DNA sequences are

¹*Feist Publ’n, Inc. v. Rural Tel. Serv. Co., Inc.*, 499 U.S. 340, 345 (1991).

copyrightable under U.S. law, based on the close analogy between human-designed DNA and computer programs² and/or the characterization of human-engineered DNA as a literary work.³

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” (17 U.S.C. § 102 (2006)).⁴ Fixation can occur in any “form, manner, or medium” (H.R. REP. NO. 94-1476, at 52 (1976)). 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 a transitory duration” (17 U.S.C. § 101 (2006)). Because DNA is composed of stable chemical nucleotides, DNA sequences easily meet this requirement. Furthermore, DNA possesses definite sequences of nucleotides that can easily be determined,⁵ copies of DNA may be synthesized routinely and in effectively unlimited quantities,⁶ and molecular DNA has been known to last for at least many thousands of years with its nucleotide sequence intact.⁷ With respect to authorship, 17 U.S.C. § 102 provides that “[c]opyright protection subsists ... in original works of authorship.” Using synthetic biological techniques, it is routine to design and construct new, human-designed DNA sequences. Because a synthetic biologist designs particular DNA sequences, and “writes,” or fixes, them when she synthesizes those sequences, she is their author. Thus, her DNA sequences should qualify as “original works of authorship.” In addition, her DNA sequences can be “perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device.”⁸

Like the English alphabet of twenty-six letters, DNA is composed of an alphabet of four nucleotide “letters”: A, T, G, and C. Literary works are defined in 17 U.S.C. § 101 as “works ... expressed in words, numbers, or other verbal or numerical symbols or indicia, regardless of the nature of the material objects ... in which they are embodied.” Nucleotides, DNA, RNA, genes, amino acids, polypeptides, and proteins certainly are “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.” Furthermore, the statement “regardless of the nature of the material objects ... in which they are embodied” certainly should include DNA.

Section 102 of the Copyright Act does not restrict eligibility for copyright protection to the seven enumerated categories. Rather, the section introduces the enumerated categories with the phrase “include[s] the following categories.” In the “Definitions” section of the Copyright Act, § 101 explains that “including ... [is] illustrative and not limitative.” The House Report accompanying the 1976 Copyright Act reinforces this broad interpretation:

The use of the word “include,” as defined in section 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

²See, e.g., Irving Kayton, Copyright in Living Genetically Engineered Works, 50 GEO. WASH. L. REV. 191 (1982); 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 (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); Christopher M. Holman, Copyright for Engineered DNA: An Idea Whose Time Has Come, West Virginia Law Review, Vol. 113, pp. 699–738 (2011); Andrew W. Torrance, Synthesizing Law for Synthetic Biology, Minnesota Journal of Law, Science & Technology, Vol. 11(2), pp. 629–665, 2010; Andrew W. Torrance, DNA Copyright, Valparaiso Law Review, Vol. 46(1), pp. 1–41 (2011).

³See, e.g., Irving Kayton, Copyright in Living Genetically Engineered Works, 50 GEO. WASH. L. REV. 191 (1982); 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 (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); Christopher M. Holman, Copyright for Engineered DNA: An Idea Whose Time Has Come, West Virginia Law Review, Vol. 113, pp. 699–738 (2011); Andrew W. Torrance, Synthesizing Law for Synthetic Biology, Minnesota Journal of Law, Science & Technology, Vol. 11(2), pp. 629–665, 2010; Andrew W. Torrance, DNA Copyright, Valparaiso Law Review, Vol. 46(1), pp. 1–41 (2011).

⁴The following argument about human-engineered DNA sequences as literary works eligible for copyright protection under the U.S. Copyright Act is adapted from; Andrew W. Torrance, Synthesizing Law for Synthetic Biology, Minnesota Journal of Law, Science & Technology, Vol. 11(2), pp. 629–665, 2010.

⁵See, e.g., F. Sanger et al., DNA Sequencing with Chain-Terminating Inhibitors, 74 PROC. NAT’L ACAD. SCI. U.S.A. 5463 (1977).

⁶See, e.g., 2 JOSEPH SAMBROOK & DAVID W. RUSSELL, MOLECULAR CLONING: A LABORATORY MANUAL 8.4 (3d ed. 2001).

⁷See, e.g., Eske Willerslev & Alan Cooper, Ancient DNA, 272 PROC. R. SOC’Y B 3, 3–5 (2005).

⁸17 U.S.C. § 102.

subject matter, but with sufficient flexibility to free the courts from rigid or outmoded concepts of the scope of particular categories.⁹

When considered in conjunction with the expansive phrase in § 102, “any tangible medium of expression, now known or later developed,” synthetic, human-designed DNA sequences fit comfortably within the category of “literary works.”

Although today the copyrightability of computer programs is non-controversial, it is important to recall that when initially confronted with that issue the Copyright Office opined that software probably was not copyrightable. Later, in 1964 the Copyright Office began issuing some provisional registrations of computer programs under its “rule of doubt.”¹⁰

In the 1970s, as an offshoot of the legislative processes leading to the Copyright Act of 1976, Congress established the National Commission on New Technological Uses of Copyrighted Works (CONTU), a body of experts tasked to study and make recommendations as to how the copyright law should respond to various technological developments, most notably the increasing significance of computer programs.¹¹ The CONTU commission was directed specifically to consider and make recommendations with respect to the question of whether, and to what extent, computer programs could be protected under current copyright law and whether copyright law should be amended to accommodate computer programs. CONTU issued its highly influential report in 1978, which concluded not only that copyright protection for computer programs was justified both in terms of legal doctrine and innovation policy, but also that computer programs in fact already were copyrightable under both the 1976 and 1909 Copyright Acts.¹² Because computer programs already were copyrightable, no amendment to the Copyright Statute was deemed necessary to extend copyright to software, although some refinements to the statute were suggested to address some unique concerns associated with the application of copyright law to software.

After the CONTU report was released, courts issued a number of decisions finding computer programs copyrightable, even highly functional programs and programs written in machine-readable form, and the Copyright Office move beyond its “rule of doubt” to acknowledge the copyrightability of computer programs. Significantly, Congress never explicitly amended the copyright statute to extend copyright protection to cover computer programs, nor did it need to, because as pointed out in the CONTU Report computer programs satisfy the requirement of authorship set forth in both the 1976 and 1909 Copyright Acts. The Copyright Office treats computer programs as literary works, one of the categories of copyrightable subject matter enumerated in the copyright act, but it is important to recognize that the list of copyrightable categories in 17 USC 102 is nonexclusive, and that regardless of whether a computer program is considered a literary work, it is copyrightable because it meets the test of “authorship fixed in a tangible medium of expression.” Human-designed DNA, which for the purpose of copyrightability is highly analogous to computer software, likewise should be considered copyrightable either as a literary work or simply because it is a human-designed work of “authorship” fixed in a tangible medium.

Because computer programs already were copyrightable under the Copyright Act, Congress saw no need to amend the statute to include computer programs in the nonexclusive list of copyrightable subject matter. However, Congress did introduce a section into the statute creating some limitations on copyright that are specific to “computer programs,” and a definition of “computer programs.” Congress’ definition of “computer program” is instructive as to why computer programs are copyrightable. Under the statute, a computer program is defined as “a set of statements or instructions to be used directly or indirectly in a computer in order to bring about a certain result.” Under general principles of copyright law, it is uncontroversial that a set of statements or instructions can be copyrightable, if the set of statements is sufficiently expressive to qualify as a work of authorship. It is also uncontroversial that a copyrightable set of instructions can be directed towards a nonhuman, such as a machine—specifically, in the case of computer programs, towards a computer. But there is nothing in copyright law that would justify treating a set of instructions directed towards a computer any differently than a set of instructions directed towards some other machine capable of receiving and acting upon the instructions, including a biological machine such as a recombinant microorganism. Not all computer programs are copyrightable, but

⁹See H.R. REP. NO. 94-1476, at 53 (1976).

¹⁰Pamela Samuelson, *Contu Revisited: The Case Against Copyright Protection for Computer Programs in Machine-Readable Form*, 1984 Duke L.J. 663, 692–93 (1984); National Commission on New Technological Uses of Copyrighted Works, Final Report at 15 (1979).

¹¹*Id.* at 693–94.

¹²National Commission on New Technological Uses of Copyrighted Works, Final Report at 16 (1979).

only those that are expressive works of authorship. However, for the same reason computer programs are copyrightable, a human-designed DNA sequence that is sufficiently expressive to satisfy the requirement of originality and authorship is copyrightable, and should be eligible for registration.

Again, it is important to emphasize that Prancer is an entirely synthetic work of authorship. This is not a DNA sequence that was derived from nature, and hence any objections that might arise when applied to the registration of a natural DNA sequence do not apply in this instance. The Prancer DNA sequence is the result of purposeful authorship, and is both an expressive literary work per se, given the creative choices that its authors made in choosing which particular nucleotides would occupy which of its sequential positions, and an expressive means of providing a set of instructions to a biological machine, just as a copyrightable computer program is an expressive means of providing a set of instructions to a computer.

Among the attributes of computer programs that render them copyrightable is the fact that there generally are a large number of potential alternative computer programs that would direct a computer program to perform essentially the same functions. Were it not for this redundancy, a computer program might not be copyrightable under the doctrine of merger.¹³ The merger doctrine limits copyright protection to a particular expression of a set of instructions to achieve a desired function, thereby permitting others to achieve essentially the same function in a computer by using an alternate means of expressing the set of instructions.

The Prancer DNA sequence clearly avoids issues of merger, because it is but one of a plethora of alternative expressive means for achieving the desired result. Prancer is a DNA sequence that provides a set of instructions for the synthesis of a protein comprising 231 amino acids linked together in a specific order. The set of instructions is coded in the standard genetic code, and is interpretable by most living biological systems.¹⁴ The encoded protein is fluorescent, which is a useful functional attribute in biotechnology. Importantly, it must be emphasized that functionality of the encoded protein in no way precludes the copyrightability of a set of instructions directing the protein's synthesis. Computer programs typically provide a set of instructions for achieving a useful function, and courts have held that this does not preclude copyright protection.¹⁵

The Prancer DNA sequence comprises 231 codons, each codon consisting of three nucleotide bases that direct a biological system to introduce an amino acid specified by the codon into a specific location in the 231 amino acid fluorescent protein.¹⁶ Most of these codons are redundant, i.e., for most of the 20 amino acids encoded by the standard genetic code, there are anywhere from two to six alternate codons specifying the same amino acid.¹⁷ As a result, there is an astronomical number of alternate coding sequences that encode the identical 231 amino acid protein.¹⁸

There also is a huge amount of potential redundancy in the amino acid sequence of the fluorescent protein itself.¹⁹ In general, the structure-function relationship in proteins is such that many amino acid substitutions could be made throughout the 231 amino acid chain without significantly altering the function of the protein. For each of these alternate protein sequences, of course, there are innumerable redundant DNA coding sequences. Thus, a copyright on the Prancer DNA sequence in no way will impede subsequent authors from creating alternate DNA sequences encoding a set of instructions that would yield substantially the same result, any more than copyright of a complex computer program would impede others from writing an alternate computer program for achieving substantially the same function. Merger simply is not an issue with this DNA sequence.

In its letter denying registration, the Copyright Office asserts that the Prancer DNA sequence does not contain the required "minimum amount of authorship." However, as noted by then-District of Columbia Circuit Court Judge (now US Supreme Court Justice) Ruth Bader Ginsburg, "a very modest degree of intellectual labor" normally is sufficient to satisfy the requirement of creativity necessary for authorship.²⁰ The Supreme Court also has emphasized that the authorship requirement is satisfied by originality and a "modicum of creativity."²¹ The Prancer DNA sequence satisfies both these criteria. The DNA sequence is original to its author, i.e., it was designed by an employee of applicant, and not derived from nature or any other source. It also is creative, in

¹³*Morrissey v. Procter & Gamble*, 379 F.2d 675, 678–79 (1st Cir. 1967) (discussing merger doctrine).

¹⁴*Genetic code*, WIKIPEDIA, http://en.wikipedia.org/wiki/Genetic_code (last visited Nov. 18, 2012).

¹⁵*Apple Computer, Inc. v. Franklin Computer Corp.*, 714 F.2d 1240, 1247–49 (3d Cir. 1983).

¹⁶*Genetic code*, WIKIPEDIA, http://en.wikipedia.org/wiki/Genetic_code (last visited Nov. 18, 2012).

¹⁷*Id.*

¹⁸Christopher M. Holman, *Protein Similarity Score: A Simplified Version of the BLAST Score as a Superior Alternative to Percent Identity for Claiming Genuses of Related Protein Sequences*, 21 Santa Clara Computer & High Tech. L.J. 55, 71 n. 70 (2004).

¹⁹BIOINFORMATICS FOR GENETICISTS Chapter 14 (Michael R. Barnes & Ian C. Gray eds., John Wiley & Sons, Ltd. 2003).

²⁰*Atari Games Corp v. Oman*, 888 F.2d 878, 882 (D.C. Cir. 1989). *See also*, *Atari Games Corp v. Oman*, 979 F.2d 242, 246 (D.C. Cir. 1992).

²¹*Feist Publ'n, Inc. v. Rural Tel. Serv. Co., Inc.*, 499 U.S. 340, 362 (1991).

that there are multiple alternative ways of encoding the same protein (or functionally equivalent proteins), and the author used selection and judgment to arrive at the Prancer DNA sequence. The Supreme Court emphasized in *Feist* that an original selection or arrangement of uncopyrightable elements is copyrightable so long as it entails a minimal degree of creativity.²²

The decision to extend copyright protection to computer programs was based in large part on the analogy between software and traditional literary works. But the analogy between computer programs and human-designed DNA is much closer. Like a computer program, a human-designed, protein-encoding DNA sequence essentially is a set of instructions directing a machine to perform certain functions in a specified order. Computer programs and human-designed DNA both can be expressed in a format directed towards a human reader, but both are directed primarily towards a machine (a recombinant microorganism can accurately be characterized as a biological machine) capable of interpreting and acting upon the set of instructions.

Another important characteristic shared by software and human-designed genetic sequences is that, for both, the cost of development greatly exceeds the cost of duplication, owing to the fact that both can serve as the template for their own reproduction. A computer program encoded in digital form can be easily and repeatedly copied, resulting in a virtually unlimited number of essentially identical copies. In the same way, DNA serves as its own template for the reproduction of the exact copies by biological DNA replication processes, both *in vitro* and *in vivo*.

Significantly, because the copies reproduced are identical to the original, and therefore can function as additional templates for further copying, software and DNA are both susceptible to viral replication. Each time a computer program is copied, it creates a new template for copying, and the iterative copying of copies can result in the production of copies at an exponential rate, mimicking the spread of a virus. In the same manner, a DNA sequence serves as the template for production of an exact copy, and each copy likewise can serve as a template for subsequent copies. Indeed, the terms “viral replication” and “computer virus” are used widely in connection with digital files and computer software, [arising] out of the remarkable propensity of viral DNA to self-replicate. Thus, the policy concerns favoring copyright protection for software apply also to human-designed DNA.

Today we are seeing an increasing convergence of the engineering of DNA and software. An example of this can be seen in the ongoing development of DNA-based computers²³ and logic gates.²⁴ The blurring of the line between genetic engineering and software engineering also can be seen in the language used by modern synthetic biologists. For example, the BioBricks Foundation, an organization established by engineers and scientists from MIT, Harvard, and University of California, San Francisco, for years has promoted the idea that a “synthetic biologist or biological engineer can, to some extent, program living organisms in the same way a computer scientist can program a computer.”²⁵ Similarly, when the CEO of Amyris Biotechnologies was asked during an interview what his synthetic biology company does, he explained that “we are really reprogramming that yeast, almost like rewriting the software inside the yeast.”²⁶

The convergence of these technologies also is evident in the incorporation of biological principles in software design. Examples include artificial intelligence and the use of neural networks in computing. One manifestation of artificial intelligence, genetic programming, applies evolutionary algorithms inspired by biological evolution to create new computer programs.²⁷ In genetic programming, functions are represented as “chromosomes,” and the main operators used in the evolutionary algorithms are “crossover” and “mutation,” concepts taken from genetics.²⁸

Meta-Genetic Programming is a proposed technique of evolving a genetic programming system using genetic programming itself. It is based upon the premise that software chromosomes, crossovers, and mutations, like

²²*Feist Publ'n, Inc. v. Rural Tel. Serv. Co., Inc.*, 499 U.S. at 348.

²³Yaakov Benenson et al., *An Autonomous Molecular Computer for Logical Control of Gene Expression*, 429 Nature 423 (2004).

²⁴Kate McAlpine, *DNA Logic Gates Herald Injectable Computers*, 2763 New Scientist 9 (2010) (“the development of DNA logic gates “brings the prospect of injectable biocomputers programmed to target diseases as they arise”).

²⁵Trichi Saukshmya and Archana Chugh, *Intellectual property rights in synthetic biology: an anti-thesis to open access to research?*, 4 Syst Synth Biol. 241 (2010). Published online 2011 February 20. doi: 10.1007/s11693-011-9067-6.

²⁶*Reprogramming yeast to make high value carbons for use in consumer goods and energy markets*, An interview with John G. Melo, CEO of Amyris, by Boonsri Dickinson, available at <http://www.biobusiness.tv/special-interview/amrs-amyris-ceo-john-g-melo-interviewed-by-boonsri-dickinson-1421/> (last visited Nov. 19, 2012).

²⁷Available at http://en.wikipedia.org/wiki/Genetic_programming (last visited Nov. 19, 2012).

²⁸*Id.*

their real-life counterparts, should be allowed to change on their own rather than being determined by human programmer.²⁹ As these trends continue, the justification for maintaining copyright protection for software while denying it for human-designed DNA becomes increasingly questionable.

Applicant urges the Copyright Office to acknowledge and embrace the convergence of computer software and engineered DNA, and to register the Prancer DNA sequence, and similarly original human-designed DNA sequences, for essentially the same reason and to the same extent that computer programs are currently eligible for copyright registration.

Respectfully submitted,
DNA2.0, Inc.
1140 O'Brien Drive, Suite A
Menlo Park, CA 94025
1-877-DNA-TOGO
www.DNA20.com

• • •

²⁹*Id.*

Supplementary Document 2: Affirmance of Refusal for Registration



United States Copyright Office

Library of Congress • 101 Independence Avenue SE • Washington DC 20559-6000
• www.copyright.gov

February 11, 2014

DNA2.0

Attn: Howard Simon
1140 O'Brien Drive, Suite A
Menlo Park, CA 94025

Correspondence ID: 1-K1L4R4
Original Correspondence ID: 1-DH5IGE

Dear Mr. Simon:

I am writing in response to your request for reconsideration of the Registration Program's refusal to register the DNA sequence entitled the "Prancer DNA Sequence." I apologize for the delay in responding, but this request was an issue of first impression for the U.S. Copyright Office and as such, was given significant consideration prior to rendering a decision. After carefully reconsidering the registration materials and the arguments contained in your request for reconsideration, the Office affirms the refusal of registration.

The Office's decision is based primarily on three principles of copyright law. First, that to be copyrightable, a claim must be based on an "original work of authorship" falling within the congressionally established categories of authorship in title 17, 17 U.S.C. §102(a) and H.R. Rep. 94-1476, 51–53. Second, copyright protection does not "extend to any idea, procedure, process, system, method of operation, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or embodied in such work." 17 U.S.C. §102(b). Third, to be copyrightable, an original work of authorship must include a sufficient quantum of copyrightable authorship. The Office finds that the Prancer DNA Sequence fails on all three of these principles.

I. DNA falls outside of the categories of copyrightable by Congress

In your argument in support of copyright protection for the Prancer DNA Sequence, you emphasize that this is a human-engineered, synthetic DNA sequence rather than a naturally-occurring DNA sequence. Were this a naturally-occurring DNA sequence that was "discovered" by your client through a process of isolation, the matter would be simply resolved by §102 as further clarified by the *Feist* decision: "one who discovers a fact is not its 'maker' or 'originator.' The discoverer merely finds and records." *Feist Publications v. Rural Telephone Service Co.*, 499 U.S. 340, 347 (1991). Your assertion that this work is created rather than discovered will, for

purposes of this review, be accepted as such. Yet, the extent to which any genetic sequence can truly be entirely human-created is unclear, particularly with respect to copyright creativity. Since the operation of DNA is dictated by laws of biology, at least some aspects of DNA sequences are controlled by those laws. Moreover, to the extent that any sequence involves the isolation of any naturally-occurring sequence, or a derivation thereof, it may not be deemed created as opposed to discovered. It is of great concern to the Office that neither this critical distinction nor the degree of creative human authorship can be established through the examination of the deposit. The inability of the Office to independently discern new creative authorship suggests that a claim in a DNA sequence may be far better suited for the realm of patent, where a heightened standard of novelty, nonobviousness, and an examination of prior art would be considered, rather than the originality standard of copyright.

However, because this claim has been submitted for copyright registration, the Office applies the principles of copyright law. Assuming *arguendo* that this sequence is not a discovery or the isolation of a naturally-occurring sequence, it nevertheless must qualify as an original work of authorship to support a claim of copyright.

The first question to address is whether this sequence qualifies as copyrightable subject matter. You argue this sequence is fixed because “DNA is composed of stable chemical nucleotides” and because “DNA possesses definite sequences of nucleotides that can easily be determined, copies known to last for at least many thousands of years with its nucleotide sequence intact.” First Request for Reconsideration, at 2. You further argue that “[u]sing synthetic biological techniques, it is routine to design and construct new, human-designed DNA sequences. Because a synthetic biologist designs particular DNA sequences, and ‘writes,’ or fixes, them when she synthesizes those sequences, she is an author.” *Id.*

The fixation of chemical nucleotide sequences designed and constructed by a biologist does not fall within any of the congressionally established categories of authorship specified in § 102(a). You argue that those categories are illustrative and not limitative according to the language of § 102(a) and the legislative history. Coupled with the 1976 Act’s expansive definition of fixation, you find support for the Copyright Act’s ability to include synthetic DNA sequences within its subject matter.

In 2012, the U.S. Copyright Office thoroughly analyzed the relationship between the § 102(a) categories and the Office’s discretion to identify new categories of authorship or to expand the scope of existing categories. In a statement of policy, the Office stated:

This passage suggests that Congress intended the statute to be flexible as to the scope of established categories, but also that Congress [] intended to retain control of the designation of entirely new categories of authorship. The legislative history goes on to state that the illustrative nature of the section 102 categories of authorship was intended to provide “sufficient flexibility to free the courts from rigid or outmoded concepts of *the scope of particular categories*.” *Id.* at 53 (emphasis added). The flexibility granted to the courts is limited to *the scope* of the categories designated by Congress in section 102(a). Congress did not delegate authority to the courts to create new categories of authorship. Congress reserved this option to itself.

If the federal courts do not have authority to establish new categories of subject matter, it necessarily follows that the Copyright Office also has no such authority in the absence of any clear delegation of authority to the Register of Copyrights.

Statement of Policy; Registration of Compilations, 77 Fed. Reg. 37,605, 37,607 (2012).

This policy statement clarified the U.S. Copyright Office’s interpretation of congressional intent based on the language of the statute together with a complete reading of the applicable legislative history. The Office found that Congress intended to avoid exhausting only its own power to create new categories of authorship:

In using the phrase “original works of authorship,” rather than “all writings of an author” now in section 4 of the statute, the committee’s purpose is to avoid exhausting the constitutional power of Congress to legislate in this field, and to eliminate the uncertainties arising from the latter phrase. Since the present statutory language is substantially the same as the empowering language of the Constitution, a recurring question has been whether the statutory and the constitutional provisions are coextensive. If so, the courts would be faced with the alternative of holding copyrightable something that Congress clearly did not intend to protect, or of holding constitutionally incapable of copyright something that Congress might one day want to protect.

H.R. Rep. 94-1476 at 51 (1976).

Congress [chose] to provide the courts with the authority to interpret the scope of existing categories, but it retained for itself the authority to create new categories of authorship in the future, as it did with architectural works. This was further clarified in the House Report:

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. In other cases, such as photographs, sound recordings, and motion pictures, statutory enactment was deemed necessary to give them full recognition as copyrightable works.

Id.

The Office finds that synthetic DNA sequences do not fit within any of the existing categories of copyrightable authorship listed in section 102(a) and are not an extension of copyrightable subject matter that Congress already intended to be protected by copyright. Even if the Office came to the opposite conclusion, based on its prior interpretation of the statute and the legislative history, the Office would not find it prudent to interpret the scope of existing categories in a wholly new manner. While the legislative history suggested that courts did have the flexibility to interpret the scope of existing categories beyond their present limits, the Office does not find similar support for its own authority. Moreover, neither the courts nor the Office [has] authority to create new categories of authorship; this prerogative resides with Congress. The Office finds a claim in synthetic DNA sequences to be a claim in a new category of copyrightable subject matter that is presently precluded from copyright protection until such time as Congress decides it should become copyrightable subject matter.

In your letter, you argue that a synthetic DNA sequence is analogous to a computer program such that it should qualify for registration as a literary work. The Office disagrees for several reasons.

You state that prior to the CONTU Commission's Final Report and Congress's subsequent amendment to the Copyright Act adding § 117 and § 101's definition of computer programs, the U.S. Copyright Office began registering certain computer programs under a "rule of doubt." However, the CONTU Report clarifies that the Office's issuance of qualified registrations were contingent upon the presence of observable authorship ("to the extent that they incorporate authorship of the programmer's expression of original ideas, as distinguished from the ideas themselves.")¹ and the deposit of human-readable copies. National Commission on New Technological Uses of Copyrighted Works, Final Report at 15 (1979).

The Copyright Office's Circular 31D contains additional conditions precedent for the registration of computer programs at that time:

- (a) The elements of assembling, selecting, arranging, editing, and literary expression that went into the compilation of the program are sufficient to constitute original authorship.
- (b) The program has been published, with the required copyright notice; that is, "copies" (i.e., reproductions of the program in a form perceptible or capable of being made perceptible to the human eye) bearing the notice have been distributed or made available to the public.
- (c) The copies deposited for registration consist of or include reproductions in a language intelligible to human beings. If the first publication was [in] a form (such as machine-readable tape) that cannot be perceived visually or read by humans, something more (such as a print-out of the entire program) must be deposited along with two complete copies of the program as first published.
- (d) An application for registration is submitted on Form A as a "book." Detailed instructions for registration are included in the application forms.
- (e) The applicant also submits a brief explanation of the way in which the program was first made available to the public, and the form in which the copies were published. This explanation is not an essential requirement in every case, but it will generally facilitate examination of the required application, copies, and fee.

See, Copyright Office Circular 31D (1967). In addition, as quoted above from the House Report, Congress suggested that computer programs were regarded as an extension of copyrightable subject matter that Congress already intended to protect, H.R. Rep. 94-1476 at 51, and were a form of literary work. Id. at 54.

¹H.R. Report at 54.

DNA sequences are fundamentally different from computer programs. As you stated, DNA sequences are fixed because “DNA is composed of stable chemical nucleotides.” DNA sequences, whether naturally occurring or synthetic, are the result of biology or biological techniques, respectively. Biological creations do not fit within any of the existing categories of authorship. Indeed, the Copyright Office Review Board upheld a similar decision in a reconsideration of a denial of registration for genetically modified plants. The fact that Congress has chosen to provide delineated patent protection for certain biological processes, such [as] plant patents for newly invented strains of asexually reproducing plants, while precluding protection for tuber-propagated plants or wild uncultivated plants provides strong justification for leaving such decisions to Congress. Moreover, the Supreme Court’s recent decision in the *Association for Molecular Pathology v. Myriad Genetics, Inc.*, also provides reason to question whether synthetic or cDNA sequences are proper subject matter for copyright, since they are eligible for patent protection. Such concerns about the potential overlap between copyright and patent protection only strengthen the Office’s conclusion that a synthetic DNA sequence does not fall within any of the existing § 102(a) categories of authorship.

Your alternative argument that the synthetic DNA sequence is analogous to a computer program because the Prancer sequence is comprised of a set of statements or instructions. You argue: “there is nothing in copyright law that would justify treating a set of instructions directed towards a computer any differently than a set of instructions directed towards some other machine capable of receiving and acting upon the instructions, including a biological machine such as a recombinant microorganism.” First Request for Reconsideration at 4.

The Office disagrees. The definition of a computer program is “a set of statements or instructions to be used directly or indirectly in a computer in order to bring about a certain result.” 17 U.S.C. § 101. Congress added this definition after the CONTU Report to address computer programs, a form of authorship that it had previously suggested fell within the scope of literary works. The 231 codons that make up the Prancer DNA Sequence are not statements or instructions that are “*used directly or indirectly in a computer in order to bring about a certain result.*” While an organism may be analogized to a machine, it clearly is not one, and as a result falls outside of the category enumerated in the statute.

Additionally, although the deposit submitted with the application for registration contains a notation of the Prancer DNA sequence comprised of specific series of four letters, the use of letters does not transform this sequence into a literary work. Every gene sequence is represented by some sequence of these four letters to identify the sequence of the chemical compounds that these letters represent. This sequence is not literary authorship, but rather a form of notating the biological sequence. Copyright does not protect DNA sequences whether naturally occurring or synthetic.

II. Copyright does not protect processes or systems

In addition to failing to fall within the scope of a congressionally-recognized category of authorship, the U.S. Copyright Office finds that this claim is precluded from copyright protection under section 102(b) which states:

(b) In no case does copyright protection for an original work of authorship extend to any idea, procedure, process, system, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or embodied in such work.

The Prancer DNA Sequence is genetic formula for a biological system. It does not describe, explain, or illustrate anything except the genetic markers that comprise this biological organism. Therefore, there is no copyrightable expression, but rather the claim simply records the formula for this biological system or process.

III. Copyright requires sufficient creative authorship

The sequence of nucleotides in a given gene is commonly represented as a list of characters comprising the letters A, C, G and T.² These letters correspond to the compounds adenine, cytosine, guanine and thymine, respectively. In the case of a synthetic gene, the specific sequence of nucleotides is the result of some person’s choices, but those choices are not made for the purpose of artistic expression. They are made to create a specific gene that produces a “particular polypeptide.” Properly understood, the nucleotide sequence of a synthetic gene,

²The letters A, C, G, and T are technically only used to represent genes sequences in DNA. Gene sequences in RNA are represented by the A, C, G, and U, where the “U” represents the nucleotide uracil.

inasmuch as it could be conceived as a form of expression at all, is a “form of expression dictated solely by functional considerations.”³ A synthetic gene’s nucleotide sequence, be it fixed in a writing or a biological cell, does not bear the “stamp” of any author.⁴ It is a mechanistically composed term that describes the gene’s physical structure in a simple and unembellished way. The sequence simply serves to identify the gene, much as any name identifies the object to which it commonly applied.

You state, “for most of the 20 amino acids encoded by the standard genetic code, there are anywhere from two to six alternate codons specifying the same amino acid.” However, even if most of the codons are redundant, that redundancy does not transform a biological sequence into an expressive work of authorship. Similarly, the mere fact that the sequence is comprised of a sequence of four letters also does not transform this biological process into a literary work. The letters are not used expressively, but rather for a functional end—to produce an encoded protein that is fluorescent, which as you state, is “a useful functional attribute in biology,” and “is interpretable by most living biological systems.” First Request for Reconsideration at 5. It is not interpretable by humans and is not used directly or indirectly in a computer to bring about a certain result. The sequence of codons is used to represent the synthesis of a protein comprising 231 amino acids that are linked together in a specific order to be used to produce a functional result in a biological organism. In the cases where letters and [symbols] are used to create a copyrightable work of authorship, they are combined into a form that is readable and conveys meaning to a human, including with the aid of a machine or device, or, to cause a certain result in a computer. The Prancer DNA Sequence does not possess any such expression beyond its functional representation.

For the reasons stated above, the U.S. Copyright Office affirms its conclusion that the Prancer DNA Sequence does not support a claim for copyright registration.

Sincerely,

A handwritten signature in dark ink, appearing to read "Robert J. Kasunic". The signature is fluid and cursive, with the first name "Robert" and last name "Kasunic" clearly distinguishable.

Robert J. Kasunic
Associate Register of Copyrights
and Director of Copyright Policy and Practices

³*Nimmer on Copyright* §2.01[B].

⁴*See Harper & Row Pub. v. Nations Entm't*, 471 U.S. 539, 547 (1985) (“The copyright is limited to those aspects of the work—termed “expression”—that display the stamp of the author’s originality.”)



United States Copyright Office

Library of Congress • 101 Independence Avenue SE • Washington DC 20559-6000
• www.copyright.gov

* 1-K1L4R4 *

Return this sheet if you request reconsideration.

How to request reconsideration:

- Send your request in writing. (It must be received in the Copyright Office not later than three months after the date on the Office's refusal letter.)
- Explain why the claim should be registered or why it was improperly refused.
- Enclose the required fee—see below.
- Address your request to:
Copyright RAC Division
P.O. Box 71380
Washington, DC 20024-1380

Note: To expedite delivery, write "Reconsideration" on the outside of the envelope. Include the Correspondence ID Number (see above) on the first page. Indicate either "First Reconsideration" or "Second Reconsideration" as appropriate on the subject line.

Notification of decision: The Copyright Office will send a written notification of its decision, including an explanation of its reasoning.

First Request for Reconsideration: The Registration and Recordation Program Office considers the first request. If it upholds the refusal, you may submit a second request.

Second Request for Reconsideration: The Copyright Office Board of Review considers the second request. The Board consists of the Register of Copyrights and the General Counsel (or their respective designees), and a third member appointed by the Register. The Board's decision constitutes final agency action.

FEES:

First Request	\$250
Additional claim in related group	\$25
Second Request	\$500
Additional claim in related group	\$25