2022

Copyright Protection for Works in the Language of Life

Nina Srejovic
Intellectual Property and Information Policy Clinic, Georgetown University Law Center,
ns1258@georgetown.edu

Washington Law Review, Forthcoming.

COPYRIGHT PROTECTION FOR WORKS IN THE LANGUAGE OF LIFE

(FORTHCOMING IN THE WASHINGTON LAW REVIEW)

NINA SREJOVIC[†]

Abstract

In 2001, the DNA Copyright Institute sought to capitalize on the fear of human cloning by offering celebrities the opportunity to use copyright to secure exclusive rights in their DNA. At the time, a Copyright Office spokesperson pointed out that a person's DNA "is not an original work of authorship." That statement is no longer self-evident. A scientist claims to have used CRISPR technology to create a pair of twin girls with human-altered DNA that may provide immunity to HIV infection and improved cognitive function. Through gene therapy, doctors can "author" changes to patients' DNA to cure disease. Scientists "edit" bacterial cell DNA to produce medicines and industrial enzymes. Researchers have "written" original DNA encoding a GIF of a running horse. Does copyright grant exclusive rights to these creations?

For decades, scholars have argued that DNA sequences, like computer programs, are copyrightable "works" encompassed by the Copyright Act's definition of "literary works." So far, the Copyright Office is unconvinced and continues to list DNA sequences and compounds as "works" that do not constitute copyrightable subject matter. This Article takes a new approach by proposing that DNA is not a "work" at all. Rather, DNA is a medium in which information is stored. In the words of the Copyright Act, DNA compounds are "copies" in which an original copyrightable work or a functional creation may be fixed. Under this framework, literature is entitled to copyright protection whether it exists as a copy printed on paper or encoded into DNA. Genetic DNA, which functions as a

---

[†] BA, Stanford University; JD, University of Michigan Law School; Clinical Teaching Fellow with the Intellectual Property and Information Policy Clinic at Georgetown University Law Center. The author would like to thank Ann Bartow, Julie Cohen, Amanda Levendowski, Christopher Morten, Madhavi Sunder, and Robin West and her SJD Colloquium for their valuable insights and assistance that contributed to this Article. This Article benefitted from presentation at the Intellectual Property Scholars Conference and the Works-in-Progress Intellectual Property Colloquium.

WORKING DRAFT
COMMENTS WELCOME

component of cellular machinery to produce useful chemicals, is entitled to no more copyright protection any other machine component. Rejecting this approach and continuing to treat DNA as a "work" rather than a "copy" has real world consequences. The recent history of copyright protection for computer programs provides a cautionary tale. Mischaracterizing DNA in the way that computer programs have been mischaracterized – as a type of "work" under the Copyright Act – could lead to the extension of exclusive copyrights to the functional DNA in living organisms in the same way that copyright protection has been extended to some functional aspects of computer programs.

TABLE OF CONTENTS

WORKING DRAFT
COMMENTS WELCOME

I. Introduction

In November 2018, Dr. He Jiankui claimed to have created the first babies born with human-edited DNA. Although He's claim has not been verified, Hank Greeley's recent book, *CRISPR People, The Science and Ethics of Editing Humans*, provides a fascinating narrative of He's experiment.[1] He and his team recruited HIV-positive couples and offered them free fertility treatments, medical care, and a stipend. According to He, he used CRISPR technology to edit the DNA of at least two fertilized eggs to grant immunity to HIV infection,[2] and these eggs resulted in the birth of a pair of genetically HIV-resistant twin girls.[3]

He's experiment raises important ethical and moral questions that have been addressed by Greeley[4] and others,[5] but this Article discusses another crucial issue exposed by recent advances in DNA technology – ownership. If He and his team were successful in altering the DNA of the fertilized eggs, did they "author" the DNA of the baby girls? If so, as authors, do copyright laws grant the exclusive right to reproduce or prepare derivative works from the girls' DNA? Following the logic of previous articles on DNA copyright, they would.[6]

---

[1] *See* HENRY T. GREELEY, CRISPR PEOPLE, THE SCIENCE AND ETHICS OF EDITING HUMANS (2020).
[2] https://www.technologyreview.com/2019/02/21/137309/the-crispr-twins-had-their-brains-altered/
[3] *See* HENRY T. GREELEY, CRISPR PEOPLE, THE SCIENCE AND ETHICS OF EDITING HUMANS (2020).
[4] HENRY T. GREELEY, CRISPR PEOPLE, THE SCIENCE AND ETHICS OF EDITING HUMANS (2020); Henry T. Greeley, *CRISPR'd babies: human germline genome editing in the 'He Jiankui affair,'* 6 J. L. AND BIO. 111 (2019). If the experiment transpired as he described, the ethical issues are more serious than those present with other DNA technology because the edits that Dr. He says he made to the babies' DNA could be passed on to future generations.
[5] *See* LeRoy Walters, Robert M. Cook-Deegan, and Eli Y. Adashi, *Governing Heritable Human Genome Editing: A Textual History and a Proposal for the Future*, forthcoming.
[6] *See, e.g.*, Irving Kayton, *Copyright in Genetic Works*, 50 GEO. WASH. L. REV. 191, 218 (1982) ("[V]irtually all original works of a genetic scientist are copyrighted automatically when he creates them."); Andrew W. Torrance, *DNA Copyright*, 46 VAL. U. L. REV. 1, 35 (2011) ("Works of genetic authorship fit within the existing framework of copyright law."); Christopher M. Holman, Charting the Contours of a Copyright Regime Optimized for Engineered Genetic Code, 69:3 OKLA. L. REV. 399, 456 (2017) ("[I]t only makes sense to move toward a copyright regime that accommodates genetic sequences. . . ."); Michael D. Murray, Post-Myriad Genetics Copyright

While the CRISPR technology that He claims to have used to edit DNA was invented in the 2010s,[7] for decades scientists have applied other techniques to construct original DNA compounds not found in nature. These techniques include splicing naturally occurring DNA compounds together to create new compounds that enable the cells of organisms to produce proteins that those cells do not produce in nature, such as human hormones,[8] improved enzymes to confront pollution,[9] cheese enzymes,[10] fuels, plastics and detergents.[11] Through gene therapy,[12] doctors can introduce beneficial alterations to patients' DNA.[13] The United States Federal Food and Drug Administration has approved several uses of human-constructed DNA to treat disease.[14] Indeed, researchers produced mRNA (a close cousin to

---

of Synthetic Biology and Living Media, 10 Okla. J.L. & Tech. 1, 30, 56 (2014) (extending the metaphor that human-constructed DNA sequences are computer programs for cells to conclude that "the entire creation of the biologist may be protected" by copyright); Devdatta Malshe, *Copyrighting DNA: An Off-Label Use,* 19 WAKE FOREST J. OF BUS. AND INTELL. PROP. L. 34, 42 (2018) (suggesting that printing the details of human-constructed DNA grants copyright protection to the DNA); Christopher M. Holman, Claes Gustafsson & Andrew W. Torrance, *Are Engineered Genetic Sequences Copyrightable?: The U.S. Copyright Office Addresses a Matter of First Impression*, 35 BIOTECH. L. REP. 103, 118 (2016) ("[T]he justification for maintaining copyright protection for software while denying it for human-designed DNA becomes increasingly questionable."); Donna Smith, Comment, *Copyright Protection for the Intellectual Property Rights to Recombinant Deoxyribonucleic Acid: A Proposal*, 19 ST. MARY'S L.J. 1083, 1096, 1106 (1988) (Recombinant DNA molecules should be copyrightable just as machine readable computer programs are.).

[7] https://science.sciencemag.org/content/337/6096/816.long

[8] NAT'L INST. OF HEALTH, HOW DID THEY MAKE INSULIN FROM RECOMBINANT DNA, https://www.nlm.nih.gov/exhibition/fromdnatobeer/exhibition-interactive/recombinant-DNA/recombinant-dna-technology-alternative.html.

[9] Lynne Peeples, *How rabbit genes could turn ordinary houseplants into pollution-eating machines*, NBC NEWS (March 5, 2019) https://www.nbcnews.com/mach/science/how-rabbit-genes-could-turn-ordinary-houseplants-pollution-eating-machines-ncna979486.

[10] C.L. Hicks, *Use of Recombinant Chymosin in the Manufacture of Cheddar and Colby Cheese*, 71 J. OF DAIRY SCI. 1127 (1988).

[11] BIO, HEALING, FUELING, FEEDING: HOW BIOTECH. IS ENRICHING YOUR LIFE (May 2010), https://www.bio.org/sites/default/files/legacy/bioorg/docs/ValueofBiotech.pdf.

[12] https://www.fda.gov/consumers/consumer-updates/what-gene-therapy-how-does-it-work?%20how%20does%20it%20work?

[13] Karen Bulaklak and Charles A. Gersbach, *The once and future gene therapy*, 11 NATURE COMMUNICATIONS 5820 (2020).

[14] https://www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products/approved-cellular-and-gene-therapy-products ("In gene therapy, scientists can do one of several things

DNA)[15] compounds which, when injected into humans, harness human cells to produce a portion of a nonhuman protein to vaccinate against Covid-19.[16]

Scientists are using both CRISPR and other DNA synthesis techniques to construct DNA for a new engineering discipline. The synthetic biology community is working to create a collection of interchangeable standard biological parts to aid in assembling engineered biological systems.[17] These "parts" are essentially a database of DNA sequences that contain the information necessary for cells to perform a standard set of biological operations. Scientists at the Massachusetts Institute of Technology (MIT) started a registry of "biological parts" to be used as standard components to construct systems or more complex parts,[18] which is now maintained by the iGEM Foundation, an organization that holds competitions for biologically engineered inventions.[19] The DNA sequences in the database may be inserted into the DNA of a cell to cause it to release a smell,[20] synthesize plastics,[21] or cause cell death.[22]

Have scientists authored these DNA compounds that produce proteins or serve other functional purposes in cells? Does copyright grant the scientists who create new functional DNA compounds exclusive rights to reproduce and distribute

---

depending on the problem that is present. They can replace a gene that causes a medical problem with one that doesn't, add genes to help the body to fight or treat disease, or turn off genes that are causing problems.")

[15] In this Article, I use the term DNA to refer to all chemical compounds that consist of a series of nucleotides, such as DNA, RNA, cDNA, oligonucleotides. While there are differences between these compounds that are relevant in the context of biochemistry, they present the same issues with regard to copyright law.

[16] *Understanding mRNA COVID-19 Vaccines,* CENTERS FOR DISEASE CONTROL AND PREVENTION (Dec. 18, 2020), https://www.cdc.gov/coronavirus/2019-ncov/vaccines/different-vaccines/mrna.html.

[17] David Singh Grewal, *Before Peer Production: Infrastructure Gaps and the Architecture of Openness in Synthetic Biology*, 20 Stan. Tech. L. Rev. 143,160 (2017).

[18] *See* Sapna Kumar and Arti Rai, *Synthetic Biology: The Intellectual Property Puzzle,* 85 TEXAS L. REV. 1745, 1745 (2007).

[19] igem, Registry of Standard Biological Parts, http://parts.igem.org

[20] igem, Registry of Standard Biological Parts, Odor, http://parts.igem.org/Odor.

[21] igem, Registry of Standard Biological Parts, Biosynthesis, http://parts.igem.org/Biosynthesis.

[22] http://parts.igem.org/Cell_death

them? Extension of copyright protection to such clearly functional DNA compounds may seem unlikely. However, persistent arguments for copyright protection for DNA due to its similarity to computer software, the current state of patent protection for DNA, and new technological developments mandate that the potential for exclusive rights granted by copyright, and their limits, be carefully considered.

Soon after the enactment of the Copyright Act of 1976, scholars began speculating about the relationship between copyright and DNA.[23] This interest has continued unabated,[24] with much legal scholarship advocating for the extension of copyright protection to DNA.[25] Scholars have concluded that by analogy to

[23] Irving Kayton, *Copyright in Genetic Works*, 50 GEO. WASH. L. REV. 191(1982); Jorge A. Goldstein, *Copyrightability of Genetic Works*, NATURE BIOTECH. 138 (Feb. 1984); Donna Smith, Comment, *Copyright Protection for the Intellectual Property Rights to Recombinant Deoxyribonucleic Acid: A Proposal*, 19 ST. MARY'S L.J. 1083 (1988); Dan L. Burk Copyrightability of Recombinant DNA Sequences, 29 Jurimetrics J. 469 (1988-89). Some have also posed the related question of copyright protection for genetically engineered organisms. *See*, Irving Kayton, *Copyright in Genetic Works*, 50 GEO. WASH. L. REV. 191, 218 (1982).\; Jorge A. Goldstein, *Copyrightability of Genetic Works*, NATURE BIOTECH. 138, 139 (Feb. 1984); For a discussion of copyright for genetically engineered organisms, *see,* Michael D. Murray, Post-Myriad Genetics Copyright of Synthetic Biology and Living Media, 10 Okla. J.L. & Tech. 1 (2014).

[24] Dan L. Burk, DNA Copyright in the Administrative State, 51 U.C. Davis Law Review (2018); Andrew Torrance, *Synthesizing Law for Synthetic Biology*, 11(2) MINN. J. L., SCI. AND TECH. 629, 642-48 (2010). Christopher M. Holman, Claes Gustafsson & Andrew W. Torrance, *Are Engineered Genetic Sequences Copyrightable?: The U.S. Copyright Office Addresses a Matter of First Impression*, 35 BIOTECH. L. REP. 103 (2016); Christopher M. Holman, Charting the Contours of a Copyright Regime Optimized for Engineered Genetic Code, 69:3 Okla. L. REV. 399, 402 (2017).

[25] Donna Comment, *Copyright Protection for the Intellectual Property Rights to Recombinant Deoxyribonucleic Acid: A Proposal*, 19 ST. MARY'S L.J. 1083, 1096, 1106 (1988) (Because recombinant DNA molecules could be viewed as "machine readable," they should be copyrightable just as machine readable computer programs.); Irving Kayton, *Copyright in Genetic Works*, 50 GEO. WASH. L. REV. 191, 201 (1982). (Engineered genetic works "are certainly analogous, if not nearly identical, to computer programs . . . [and b]ecause of this similarity . . . should be copyrightable."); Christopher M. Holman, Claes Gustafsson & Andrew W. Torrance, *Are Engineered Genetic Sequences Copyrightable?: The U.S. Copyright Office Addresses a Matter of First Impression*, 35 BIOTECH. L. REP. 103, 118 (2016) ("[T]he justification for maintaining copyright protection for software while denying it for human-designed DNA becomes increasingly questionable."). Dan L. Burk and Iver P. Cooper represent the dissenting opinion. *See*

copyright in computer programs, copyright should protect "human-designed DNA,"[26] "DNA code,"[27] "genetic works,"[28] or "recombinant DNA."[29]

Few would argue that the model of copyright protection for computer software is something to emulate. Since registration of the first computer program in the Copyright Office, both courts and commentators have shown significant ambivalence concerning copyright protection for computer programs. In the decades since the Commission on New and Technological Uses[30] (CONTU) declared that copyright protects computer programs but not the electro-mechanical functioning of a machine, courts have tied themselves in knots in their attempts to distinguish between the two. No fewer than four different tests have been devised in the various federal circuits to separate the copyrightable "expression" of computer programs from their uncopyrightable function.[31] As stated by the Second

---

Dan L Burk, *Copyrightability of Recombinant DNA Sequences*, 29 JURIMETRICS J. 469 (1989); 2 BIOTECHNOLOGY AND THE LAW § 11.02 (rev. ed. 2000.

[26] Christopher M. Holman, Claes Gustafsson & Andrew W. Torrance, *Are Engineered Genetic Sequences Copyrightable?: The U.S. Copyright Office Addresses a Matter of First Impression*, 35 BIOTECH. L. REP. 103, 118 (2016) ("[T]he justification for maintaining copyright protection for software while denying it for human-designed DNA becomes increasingly questionable.").

[27] Devdatta Malshe, *Copyrighting DNA: An Off-Label Use*, 19 WAKE FOREST J. OF BUS. AND INTELL. PROP. L. 34, 42 (2018) (questioning the logic of protection for "computer code" while denying it to "DNA code").

[28] Irving Kayton, *Copyright in Genetic Works*, 50 GEO. WASH. L. REV. 191, 201 (1982). (Engineered genetic works "are certainly analogous, if not nearly identical, to computer programs . . . [and b]ecause of this similarity . . . should be copyrightable.").

[29] ; Comment, *Copyright Protection for the Intellectual Property Rights to Recombinant Deoxyribonucleic Acid: A Proposal*, 19 ST. MARY'S L.J. 1083, 1096, 1106 (1988) (Because recombinant DNA molecules could be viewed as "machine readable," they should be copyrightable just as machine readable computer programs.).

[30] For a description of the Commission and its discussions concerning software, *see*, Pamela Samuelson, *Contu Revisited: The Case Against Copyright Protection For Computer Programs in Machine-Readable Form*, 1984 DUKE L.J. 663.

[31] *See*, Pamela Samuelson, *Functionality and Expression in Computer Programs: Refining the Tests for Software Copyright Infringement* 31 BERKELEY TECH. L. J. 1215 (2017) (evaluating four tests used by courts in determining copyright infringement of computer programs: 1) an approach treating all structure, sequence, and organization of programs as protectable expression as long as there are multiple ways to perform a program function; 2) an approach applying an abstraction-filtration-comparison test; 3) an approach focused on whether elements are processes or methods

Circuit court in *Computer Assocs. Int'l v. Altai, Inc.*, "[t]hus far, many of the decisions in this area reflect the courts' attempt to fit the proverbial square peg in a round hole." [32] As a result, court decisions regarding the scope of copyright protection in computer technology are highly unpredictable. In *Oracle v. Google*, the Supreme Court could do no better than to start its reasoning with the *assumption* that the program copied was entitled to copyright protection. In addition, courts have upheld exclusive rights in functional technology under copyright without the novelty required by the patent system or the limited duration of rights granted by patent. Following the lead of copyright protection for computer technology in the context of DNA technology would lead to similar uncertainty and exclusive rights in the functional aspects of cellular machinery.

Recent experience with patent protection for DNA also points to the importance of carefully considering copyright protection for DNA technology. Scholars have previously recognized the risks of extending broad patent rights to human-constructed DNA that operates in cells of organisms engineered to have new abilities. Sapna Kumar and Arti Rai have noted that foundational patents and patent thickets in the context of DNA have the potential to stifle innovation to the extent they cover standards that synthetic biologists seek to establish.[33] The goals of standardization may be thwarted,[34] and subsequent research may suffer from lack of access.[35] While, as Kumar and Rai correctly recognize, intellectual property

---

of operation excluded from copyright protection; and 4) an approach concentrated on determining whether program ideas or functions have merged with program expression).

[32] Computer Assocs. Int'l v. Altai, Inc., 982 F.2d 693, 712 (2d Cir. 1992)

[33] Sapna Kumar and Arti Rai, *Synthetic Biology: The Intellectual Property Puzzle*, 85 TEXAS L. REV. 1745, 1747 (2007).

[34] Sapna Kumar and Arti Rai, *Synthetic Biology: The Intellectual Property Puzzle*, 85 TEXAS L. REV. 1745, 1747-62 (2007).

[35] Recognizing that in the context of synthetic biology, "the ability to invoke copyright [was] by no means clear," Kumar and Rai advocate for a "parallel unpatented space." *See* Sapna Kumar and Arti Rai, *Synthetic Biology: The Intellectual Property Puzzle*, 85 TEXAS L. REV. 1745, 1764, 1768 (2007). Similarly, discussions of the obstacles to establishing an "open source" database of these DNA sequences have focused on patent, rather than copyright, protection for DNA. *See* David Singh Grewal, Before Peer Production: Infrastructure Gaps and the Architecture of openness in Synthetic Biology, 20 Stan. Tech. L. Rev. 143, 178-79 (2017). *See also*, Ethan R. Fitzpatick, Open Source Synthetic biology: Problems and Solutions, 43 Seton Hall L. Rev. 1363, 1378 (2013). Interestingly, the BioBricks Foundation, which seeks to use contract law to establish a public domain of DNA sequences to promote open development in this technology, requires nonassertion

rights may sometimes be necessary to create a commons for use by multiple parties,[36] if the experience of patent rights in DNA is any guide, it seems more likely that any exclusive copyrights granted in DNA will be closely guarded. Recent scholarship has also predicted that the Supreme Court's decision in *Association for Molecular Pathology v. Myriad Genetics*— invalidating patent claims for some DNA compounds— will only intensify the interest in copyright protection for DNA because those who have previously relied on patents to gain exclusive rights in DNA sequences may turn to copyright instead.[37]

Recent technological advances make clear that categorical statements about the copyrightability of DNA are not sufficient to address the different types of information stored in DNA. The Copyright Office's Compendium states that "DNA sequences and other genetic, biological, or chemical substances or compounds" as a rule do not constitute copyrightable subject matter,[38] The Compendium provides no support for such a blanket statement other than conclusively stating that such sequences are examples of the ideas, procedures, processes, systems, methods of operation, concepts, principles, or discoveries that are excluded from copyright protection under section 102(b) of the Copyright Act.[39] But scientists are

---

of all intellectual property rights. *See* BIOBRICKS FOUNDATION, THE OPEN MATERIAL TRANSFER AGREEMENT, https://biobricks.org/open-material-transfer-agreement/.

[36] Sapna Kumar and Arti Rai, *Synthetic Biology: The Intellectual Property Puzzle*, 85 TEXAS L. REV. 1745, 1748 (2007) ("Yet many of the techniques of open source *require* property rights so that future users and third parties will be bound by the terms of the license.").

[37] Devdatta Malshe, *Copyrighting DNA: An Off-Label Use*, 19 WAKE FOREST J. OF BUS. AND INTELL. PROP. L. 34, 37 (2018) (Predicting that the resulting action from the Supreme Court's *Myriad* decision "is now going to be a scramble to get man-made DNA copyright protection."); Andrew W. Torrance and Linda J. Kahl, *Bringing Standards to Life: Synthetic Biology Standards and Intellectual Property*, SANTA CLARA HIGH TECH. L. REV. 227 (2014) ("now that natural-source DNA molecules have lost their eligibility for patent protection, copyright stands ready to provide an existing alternative form of protection."); This Author, when in private practice, personally experienced this increased interest in copyright protection for DNA.

[38] U.S. COPYRIGHT OFF., COMPENDIUM OF U.S. COPYRIGHT OFF. PRACTICES § 313.3(A) (3d ed, 2021).

[39] A letter from Robert J. Kasunic, Associate Register of Copyrights and Director of Copyright Policy and Practices in response to a request by Christopher Holman and Andrew Torrance and Dr. Claes Gustafsson for reconsideration of a refusal to register a specific human-constructed DNA compound labeled the "Prancer DNA Sequence" provides a window into the reasoning behind the statement in the Compendium. The letter states that 1) the Prancer DNA Sequence is

experimenting with DNA compounds as storage devices for information seemingly unanticipated by the Copyright Office. The Office of the Director of National Intelligence has established the Molecular Information Storage (MIST) research program to test DNA, among other chemical compounds, as a storage technology "with reduced physical footprint, power and cost requirements relative to conventional storage [i.e., digital] technologies."[40] Private sector companies and academic researchers also are exploring the use of DNA for dense and durable information storage, both inside[41] and outside[42] of living cells. These innovations use DNA in ways that are completely unrelated to what most people consider to be DNA's function in nature, ensuring the development, survival, and reproduction of a living organism. Researchers have encoded a video from the Warsaw Ghetto Archives in DNA and embedded it in a pair of eyeglasses with standard transparent lenses.[43] They constructed a 3-D printed bunny-shaped trinket called the Stanford

---

not "within the congressionally established categories of authorship in title 17;" 2) the Prancer DNA Sequence does "not include a sufficient quantum of copyrightable authorship," and 3) copyright protection is precluded for the Prancer DNA Sequence because protection does not extend to any idea, procedure, process, system, method of operation, principle, or discovery. February 11, 2014 letter from Robert J. Kasunic to Howard Simon reprinted as Supplementary Document 2: Affirmance of Refusal for Registration in Christopher M. Holman, Claes Gustafsson & Andrew W. Torrance, *Are Engineered Genetic Sequences Copyrightable?: The U.S. Copyright Office Addresses a Matter of First Impression*, 35 BIOTECH. L. REP. 103 (2016).

[40] OFF. OF THE DIR. OF NAT'L INTELL., MOLECULAR INFO. STORAGE (MIST), (last visited Feb. 9, 2021), www.iarpa.gov/index.php/research-programs/mist/mist-baa.

[41] Seth L. Shipman, et al., *CRISPR-Cas encoding of a digital movie into the genome population of living bacteria*, 547 NATURE 345 (2017); Ed Yong, *Scientist Can Use CRISPR to Store Images and Movies in Bacteria*, THE ATL. (July 12, 2017), https://www.theatlantic.com/science/archive/2017/07/scientists-can-use-crispr-to-store-images-and-movies-in-bacteria/533400/.

[42] Mike Brunker, *Microsoft and Univ. of Wash. Researchers set record for DNA storage*, THE AI BLOG (July 7, 2016), https://blogs.microsoft.com/ai/synthetic-dna-storage-milestone/#sm.0000k81a37qr6dijzdl15reujpheo.

[43] Julian Koch, Silvan Gantenbein, Kunal Masania, Wendelin J. Stark, Yaniv Erlich, Robert N. Grass, *A DNA storage architecture to create materials embedded with memory*, 38 NATURE BIOTECH. 39, 42 (2020). https://www.nature.com/articles/s41587-019-0356-z.epdf?sharing_token=c2Pos7WwuJ9MtMGj-qdFQNRgN0jAjWel9jnR3ZoTv0M4Woj1cE3OBfuw5I5lxno_c7GoY2-6n89GH-ivEpAEquXjCDjHA8AZdMxio1_l4363ezz9mt81f8Ux0-ThMicOcJ3jN17Y29Zjoyaxwr_igBrd3adSox6_-oH3cNq6DJ1ULcp4ByGA1x5klZvn7uokkqRdDbeKAWtoswVNzrPPjDH9ZvIMjCntNYKs1wFv

Bunny that contains, in the material used to make it, DNA compounds storing the information necessary to 3-D print another copy of the bunny.[44] Some predict that as the technology advances, the process will soon be used to create a "DNA of things."[45] Applications for the process may be mundane, such as a car bumper having the instructions necessary to 3-D print a replacement in the case of damage, or, as Drew Endy has mused, the uses may be sublime:

> *Imagine a societal norm in which every object must encode the instructions for making the object. Given the incredible information density of DNA data storage, such information could, in some commonplace objects such as refrigerators, also include a fully unabridged guide to rebuilding all of civilization.*[46]

No previous scholarship has considered whether copyright grants exclusive rights

---

1aqEveze5ycUK_kclvwA58FZVaUZfx68IiSptq24UNUNJR2zrvMR0Vwz_5wNAwNHnqY%3D &tracking_referrer=spectrum.ieee.org

[44] Julian Koch, Silvan Gantenbein, Kunal Masania, Wendelin J. Stark, Yaniv Erlich, Robert N. Grass, *A DNA storage architecture to create materials embedded with memory*, 38 NATURE BIOTECH. 39, 40 (2020). https://www.nature.com/articles/s41587-019-0356-z.epdf?sharing_token=c2Pos7WwuJ9MtMGj-qdFQNRgN0jAjWel9jnR3ZoTv0M4Woj1cE3OBfuw5I5lxno_c7GoY2-6n89GH-ivEpAEquXjCDjHA8AZdMxio1_l4363ezz9mt81f8Ux0-ThMicOcJ3jN17Y29Zjoyaxwr_igBrd3adSox6_-oH3cNq6DJ1ULcp4ByGA1x5klZvn7uokkqRdDbeKAWtoswVNzrPPjDH9ZvIMjCntNYKs1wFv1aqEveze5ycUK_kclvwA58FZVaUZfx68IiSptq24UNUNJR2zrvMR0Vwz_5wNAwNHnqY%3D &tracking_referrer=spectrum.ieee.org

[45] Julian Koch, Silvan Gantenbein, Kunal Masania, Wendelin J. Stark, Yaniv Erlich, Robert N. Grass, *A DNA storage architecture to create materials embedded with memory*, 38 NATURE BIOTECH. 39, 39 (2020). https://www.nature.com/articles/s41587-019-0356-z.epdf?sharing_token=c2Pos7WwuJ9MtMGj-qdFQNRgN0jAjWel9jnR3ZoTv0M4Woj1cE3OBfuw5I5lxno_c7GoY2-6n89GH-ivEpAEquXjCDjHA8AZdMxio1_l4363ezz9mt81f8Ux0-ThMicOcJ3jN17Y29Zjoyaxwr_igBrd3adSox6_-oH3cNq6DJ1ULcp4ByGA1x5klZvn7uokkqRdDbeKAWtoswVNzrPPjDH9ZvIMjCntNYKs1wFv1aqEveze5ycUK_kclvwA58FZVaUZfx68IiSptq24UNUNJR2zrvMR0Vwz_5wNAwNHnqY%3D &tracking_referrer=spectrum.ieee.org

[46] Emily Waltz, *With DNA Data Storage, 3D-Printed Bunnies Carry Their Own Blueprints*, HUMAN OS BLOG (Dec. 9, 2019), https://spectrum.ieee.org/the-human-os/biomedical/devices/dna-of-things.

in DNA compounds used in this way for data storage.

This Article analyzes how the Copyright Act intersects with DNA technology without relying on an analogy to the troubled state of copyright protection for computer technology or on the unsupported blanket exclusion enunciated by the Copyright Office. As this Article will demonstrate, as long as DNA technology is correctly situated within the Copyright Act, copyright does not grant exclusive rights to functional, and more specifically genetic, DNA. Despite the history of copyright protection for computer programs, there is no need to divine the intent of the drafters of the Copyright Act to determine whether DNA is a new type of copyrightable work encompassed by the statute. DNA is not a "work" at all, but rather, a medium in which information is stored. Granting copyright protection for literature or music stored in DNA compounds should not lead to categorical copyright protection for DNA any more than copyright protection for literature written on paper or music recorded on magnetic tape should lead to copyright protection for paper or tape. Similarly, cellular processes or functional proteins stored in DNA should not be protected by copyright simply because they are stored in DNA. Once DNA is recognized as the physical object in which information is stored, copyright protection for that information can be assessed for copyright just as information stored in any other form. In addition, copyright protection for verbal representations of DNA compounds as literary works is separate from copyright protection for the DNA compound itself and is subject to the same limitations as copyright protection for any other literary work. As long as this framework is maintained, it will be clear that even if copyright subsists in copyrightable works stored in DNA compounds and in verbal representations of some DNA, copyright protection does not extend to functional genetic[47] DNA.

This Article contains three Parts beyond this Introduction. Part II of this article will discuss DNA's role as an information technology. Part III will address, in turn, copyright protection for 1) expressive information and 2) functional information stored in DNA compounds. Part IV will discuss the somewhat thorny metaphysics of the scope of copyright protection for verbal representations of DNA

[47] I use the term "genetic" to describe DNA compounds, or portions of such compounds, that are capable of operating in cells to produce proteins. They may be either naturally occurring or human -constructed DNA compounds.

compounds as literary works.

II. DNA As Information Technology

While the motivation to construct new DNA compounds may vary from producing better humans to manufacturing proteins to storing our vast stores of digital data, each application of DNA technology takes advantage of a core attribute of DNA within biological systems: the ability of DNA to store, transfer and retrieve information.[48] As Arti Rai recognized more than 20 years ago, "[a]lthough DNA is, obviously enough, a chemical compound, it is more fundamentally a carrier of information."[49] More recently, discussions of the various court decisions in the dispute between the Association for Molecular Pathology and Myriad Genetics regarding the patentability of DNA similarly reference the information-carrying function of DNA.[50]

In the case of naturally occurring genetic DNA in living cells, DNA carries the information necessary to produce all the proteins required for survival of the organism. The information in DNA functions as the cell's "operating system" much as operating system programs function within computers. Living cells are protein

---

[48] Although Dan Burk correctly points out that other biological compounds participate in the information transfer system in a cell and references to the uniqueness of DNA among chemical compounds suffer from the notion of the "DNA mystique," his analogy of DNA to a cog in Babbage's "famous 'difference engine,' the conceptual precursor to modern computing," demonstrates that while other molecules may store information, that fact does not diminish DNA's role as a form of information technology. Dan L. Burk, *The Problem of Process in Biotechnology*, 43 HOUS. L. REV. 561, 583 (2006).

[49] Arti K. Rai, *Intell. Prop. Rights in Biotechnology: Addressing New Tech.*, 34 WAKE FOREST L. REV. 827, 836 (1999) ("The CAFC's failure to recognize DNA-based technologies as involving information first and foremost reveals its inability to adjust existing paradigms to address new technology."); *See also* Rebecca Eisenberg, *Molecules vs. Information: Should Patents Protect Both?*, 8 B.U. J. SCI . & TECH. L. 190 (2002).

[50] Christopher M. Holman, *Developments in Synthetic Biology Are Altering the IP Imperatives of Biotechnology*, 17 VAN. J. ENT. & TECH. L. 385, 461 (2015). *See also*, Dan L. Burk, *The Curious Incident of the Supreme Court in Myriad Genetics*, 90 NOTRE DAME L. REV. 505 (2014); Ethan R. Fitzpatrick, *Open Source Synthetic Biology: Problems and Solutions*, 43 SETON HALL L. REV. 1363, 1385-1387 (2013) (discussing the emphasis placed on the information-carrying qualities of DNA by the Supreme Court in *Myriad*).

producing machines. Rather than taking in digital input and producing digital output as computers do, they take in chemical input and produce chemical output. A chemical enzyme acts as the input to the cellular "computer" which initiates the process of producing a protein. Just as the information in a computer's operating system operates to produce a different output depending on the input, the information in a cell's DNA operates to produce a different protein depending on the enzyme introduced.

The information necessary to produce the proteins that a cell produces is stored in DNA just as the information necessary to produce the digital output that a computer produces is stored in computer software.[51] Information is stored in DNA as four different chemical subunits. The chemical subunits are called nucleotides and are arranged linearly along each of the two helical strands that make up the DNA compound. The order of these four chemical subunits in DNA can encode any type of information in DNA that the order of two electronic[52] states can encode in software. In written descriptions of DNA, the four different types of nucleotides in DNA are usually referred to by the letters A, G, C, and T just as in computer code the two different electronic states are referred to as 0 and 1.

The DNA Fact Sheet from the National Human Genome Research Institute provides a simplified example.[53] The order of nucleotides in the DNA of a person

---

[51] Indeed, as both technologies continue to evolve, the relationship between software and DNA becomes less of an analogy and more of a convergence. *See*, Douglas Carmean, *et al*., *DNA Storage and Hybrid Molecular-Electronic Computing*, 107 PROC. OF THE IEEE 63, 65-7 (2019), https://ieeexplore.ieee.org/abstract/document/8556046 (proposing hybrid molecular-electronic systems).
Luis Ceze, *et al*., *Molecular digital data storage using DNA*, 20 NATURE REV. GENETICS 456, 456 (2019) (referring to "the growing intersection of computer systems and biotechnology").

[52] Information can be stored in computer software as electromagnetic, optical, or silicon-based "on" and "off" states, but I use "electronic" as an example and for readability.

[53] *Deoxyribonucleic Acid (DNA) Fact Sheet*, NATIONAL HUMAN GENOME RESEARCH INSTITUTE (last visited Feb. 9, 2021), https://www.genome.gov/about-genomics/fact-sheets/Deoxyribonucleic-Acid-Fact-Sheet. The exact wording of the example from the DNA Fact Sheet from the National Human Genome Research Institute is problematic in the context of copyright law. The fact sheet states that the information contained in the sequence of A, T, C, and G nucleotides that are present in a DNA compound "determines what biological instructions are contained in a strand of DNA. For example, the sequence ATCGTT might instruct for blue eyes, while ATCGCT might instruct

determines eye color. The cells of people with the DNA nucleotide sequence ATCGTT in their DNA might produce the proteins that result in them having blue eyes while the cells of people with the DNA nucleotide sequence ATCGCT might produce the proteins that result in them having brown eyes.[54]

Scientists have harnessed this information-carrying quality of DNA to construct cells that manufacture desired proteins. The cells contain DNA with human-authored nucleotide sequences that function as operating systems not found in nature. Thus, altered bacterial cells operate to produce human growth hormone.[55] Altered fish cells operate to produce proteins that fluoresce when exposed to artificial light.[56] Scientists also construct cellular operating systems to produce therapeutic[57] and industrial proteins.[58]

Most of the DNA nucleotide sequences for these human-constructed cellular operating systems are authored by splicing together shorter sequences that naturally occur in living organisms.[59] However, recent advances in the ability to "write" completely novel DNA compounds has enabled researchers to envision

---

for brown." As I will explain below, the information stored in DNA should not be considered "instructions" in determining copyrightability.

[54] DNA performs this role by serving as a template in a two-step process in which DNA is "transcribed" into RNA, which is then "translated" into protein. Proteins are constructed in cells from a group of 20 different amino acids.[54] A sequence of three nucleotides in a DNA compound contains the information necessary for a cell to add a single amino acid to a protein synthesized by the cell. For example, the nucleotides CTA in sequence in a strand of DNA will add the amino acid leucine to a protein constructed by the cell. For a more detailed description of how DNA operates within cells, *see*, Assoc. for Molecular Pathology v. U.S. Pat. and Trademark Off., 702 F. Supp. 2d 181, 193-200 (S.D.N.Y. 2010), aff'd in part & rev'd in part, 689 F.3d 1303 (Fed. Cir. 2012), aff'd in part & rev'd in part, 569 U.S. 576 (2013).

[55] David V. Goeddel, et al., *Direct expression in Escherichia coli of a DNA sequence coding for human growth hormone*, 281 NATURE 544 (1979).

[56] Leslie Pray, *Recombinant DNA technology and transgenic animals*, 1 NATURE EDUCATION 51 (2008).

[57] NAT'L INST. OF HEALTH, HOW DID THEY MAKE INSULIN FROM RECOMBINANT DNA, https://www.nlm.nih.gov/exhibition/fromdnatobeer/exhibition-interactive/recombinant-DNA/recombinant-dna-technology-alternative.html.

[58] BIO, HEALING, FUELING, FEEDING: HOW BIOTECH. IS ENRICHING YOUR LIFE (May 2010), https://www.bio.org/sites/default/files/legacy/bioorg/docs/ValueofBiotech.pdf.

[59] Such "recombinant" DNA sequences have been used to construct Escherichia coli bacteria that produce a human protein or to cause fish to fluoresce when exposed to artificial light. *See*

DNA as information technology to store, retrieve and send information unrelated to protein synthesis inside[60] and outside[61] of living cells. Scientists can now imagine a sequence of DNA nucleotides, type the sequence into an online form, and receive in the mail within a matter of days a vial containing DNA compounds of that sequence. "With simple chemistry, strings of A, T, C and G nucleotides can be created in any desired order, one [nucleotide] after another, snapping together in a similar fashion to Lego pieces."[62] Automated DNA "printers" assemble DNA compounds of the desired sequence[63] by taking information from a computer database and controlling the valves of the printer to assemble a DNA compound nucleotide by nucleotide.[64]

DNA sequencing and synthesis technology has illuminated DNA's role as an information technology. Because scientists are not limited to splicing together naturally occurring DNA, DNA compounds, like digital technology, can now be constructed and used to store nearly any type of information, including literature, motion pictures, music, and even computer programs.[65] A manuscript authored by

---

[60] Ed Yong, *Scientist Can Use CRISPR to Store Images and Movies in Bacteria*, THE ATL. (July 12, 2017), https://www.theatlantic.com/science/archive/2017/07/scientists-can-use-crispr-to-store-images-and-movies-in-bacteria/533400/.

[61] Julian Koch, Silvan Gantenbein, Kunal Masania, Wendelin J. Stark, Yaniv Erlich, Robert N. Grass, *A DNA storage architecture to create materials embedded with memory*, 38 NATURE BIOTECH. 39 (2020). https://www.nature.com/articles/s41587-019-0356-z.epdf?sharing_token=c2Pos7WwuJ9MtMGj-qdFQNRgN0jAjWel9jnR3ZoTv0M4Woj1cE3OBfuw5I5lxno_c7GoY2-6n89GH-ivEpAEquXjCDjHA8AZdMxio1_l4363ezz9mt81f8Ux0-ThMicOcJ3jN17Y29Zjoyaxwr_igBrd3adSox6_-oH3cNq6DJ1ULcp4ByGA1x5klZvn7uokkqRdDbeKAWtoswVNzrPPjDH9ZvIMjCntNYKs1wFv1aqEveze5ycUK_kclvwA58FZVaUZfx68IiSptq24UNUNJR2zrvMR0Vwz_5wNAwNHnqY%3D&tracking_referrer=spectrum.ieee.org

[62] Jerry T., *How Oligos Changed the World*, TWIST BIOSCIENCE (Dec. 12, 2017), https://twistbioscience.com/company/blog/oligos-changed-world.

[63] CODEX DNA, WELCOME TO THE FUTURE (last visited Feb. 10, 2021), https://codexdna.com/pages/bioxp-3200-system

[64] Drew 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.

[65] *See, e.g.*, Nick Goldman, *et al.*, *Towards practical, high-capacity, low-maintenance information storage in synthesized DNA*, 494 NATURE 77, 77 (2013) (sonnets, a scientific paper, a color photographs, a speech); Mike Brunker, *Microsoft and Univ. of Wash. Researchers set record for*

the scientist George Church has been stored in DNA by converting the two electronic states of a digitally stored version into the A, T, C and G nucleotides of a DNA compound. The A and C nucleotides took the place of one electronic state and the T and G nucleotides took the place of the other electronic state.[66] A sonnet,[67] a motion picture,[68] photographs,[69] and the whole of Wikipedia[70] have been stored in DNA. A team at the University of Washington stored iconic musical performances from the Montreux Jazz Festival, the top 100 books of Project Gutenberg, the Universal Declaration of Human Rights in 100 languages, and the non-profit Crop Trust's entire seed database in DNA.[71] The university currently displays a portrait of Rosalind Franklin, the scientist who first discovered the helical structure of DNA, constructed by collaging approximately 50,000 photographs

DNA storage, THE AI BLOG (July 7, 2016), https://blogs.microsoft.com/ai/synthetic-dna-storage-milestone/#sm.0000k81a37qr6dijzdl15reujpheo (high definition video, the Universal Declaration of Human Rights, books, a seed database); Susan Young Rojahn, *An Entire Book Written in DNA*, MIT TECHNOLOGY REVIEW (Aug. 16, 2012), https://www.technologyreview.com/2012/08/16/184447/an-entire-book-written-in-dna/ (JavaScript program).

[66] Luis Ceze, et al., *Molecular digital data storage using DNA*, 20 NATURE REV. GENETICS 456, 459 fig. 3 (2019); Susan Young Rojahn, *An Entire Book Written in DNA*, MIT TECHNOLOGY REVIEW (Aug. 16, 2012), https://www.technologyreview.com/2012/08/16/184447/an-entire-book-written-in-dna/.

[67] Nick Goldman, *et al*., *Towards practical, high-capacity, low-maintenance information storage in synthesized DNA*, 494 NATURE 77, 77 (2013).

[68] Ed Yong, *Scientist Can Use CRISPR to Store Images and Movies in Bacteria*, THE ATL. (July 12, 2017), https://www.theatlantic.com/science/archive/2017/07/scientists-can-use-crispr-to-store-images-and-movies-in-bacteria/533400/.

[69] Nick Goldman, *et al*., *Towards practical, high-capacity, low-maintenance information storage in synthesized DNA*, 494 NATURE 77, 77 (2013).

[70] Chris Mellor, *Catalog claims DNA data storage is economically feasible for the first time*, BLOCKS & FILES (March 18, 2020), https://blocksandfiles.com/2020/03/18/catalog-cdna-data-storage-economically-feasible/.

[71] Mike Brunker, *Microsoft and Univ. of Wash. Researchers set record for DNA storage*, THE AI BLOG (July 7, 2016), https://blogs.microsoft.com/ai/synthetic-dna-storage-milestone/#sm.0000k81a37qr6dijzdl15reujpheo; #*MemoriesInDNA portrait project blends DNA technology and art to memorialize pioneering scientist Rosalind Franklin*, ALLEN SCHOOL NEWS, U. OF WASH. (Feb. 24, 2020), https://news.cs.washington.edu/2020/02/24/memoriesindna-portrait-project-blends-dna-technology-and-art-to-memorialize-pioneering-scientist-rosalind-franklin/.

collected from the public and stored as synthetic DNA.[72] The work of art is coated with ink mixed with the DNA in which the photographs are stored. As explained by one of the researchers involved in the project, if you were to scrape a bit of the portrait off, with the right equipment you could retrieve the data and convert the DNA compounds into digital data and then recreate the photographs themselves.[73]

Although DNA storage technology is still in its infancy, limited to experimental as opposed to commercial uses, it introduces a new impetus to copyright protection for DNA and a new reason to guard against its overextension.[74] The ability of DNA to store all the information necessary to produce a copy of a photograph, a figurine, a sound recording or a literary work in nearly any physical form serves to illustrate that DNA, like videocassettes and computer software, is, in fact, simply the latest available information technology. Indeed, digital data storage and DNA data storage are becoming interchangeable. A computer operating system has been stored in a DNA compound.[75] The sequence of nucleotides in a DNA compound are routinely stored as digital computer data.[76] Someday DNA data storage may supplant digital data storage. Because DNA data storage can be

---

[72] #*MemoriesInDNA portrait project blends DNA technology and art to memorialize pioneering scientist Rosalind Franklin*, ALLEN SCHOOL NEWS, U. OF WASH. (Feb. 24, 2020), https://news.cs.washington.edu/2020/02/24/memoriesindna-portrait-project-blends-dna-technology-and-art-to-memorialize-pioneering-scientist-rosalind-franklin/.

[73] #*MemoriesInDNA portrait project blends DNA technology and art to memorialize pioneering scientist Rosalind Franklin*, ALLEN SCHOOL NEWS, U. OF WASH. (Feb. 24, 2020), https://news.cs.washington.edu/2020/02/24/memoriesindna-portrait-project-blends-dna-technology-and-art-to-memorialize-pioneering-scientist-rosalind-franklin/.

[74] DNA storage technology may soon strike fear in the hearts of the owners of copyright in those works. As Jessica Litman has recounted with respect to earlier technologies that made it easier for the public to make unauthorized copies of copyrighted works, copyright holders may even fight the development of such a technology or at least put limits on its functionality. *See* JESSICA LITMAN, DIGITAL COPYRIGHT 177 (2006) (Describing efforts to prohibit the sale of videocassette recorders, prohibit the rental of records or computer software, require that recording devices be technologically equipped to prevent serial copying, and prohibit circumvention of technological protection measures controlling access to copyrightable works.).

[75] Yaniv Erlick, Dina Zielinski, *DNA Fountain enables a robust and efficient storage architecture*, 355 SCIENCE 950, 950 (2017).

[76] For example, *see*, 37 C.F.R. § 1.821(c) (2019) and 37 C.F.R. § 1.821(f) (2019) (requiring nucleotide sequences submitted to the United States Patent and Trademark Office to be submitted on paper or compact disc and in computer readable form.).

much more efficient than data storage in software,[77] DNA storage is not constrained to any particular shape,[78] and DNA is more stable than software,[79] researchers are looking to DNA as a solution to close the gap between the amount of data we produce and our capacity to store it.[80] These technological advances also make it more tempting than ever for courts to equate DNA technology with computer technology. Following the analogy to computer programs, copyright protection could be granted to human-constructed functional DNA compounds, in human or other cells, just as copyright protection has been granted by courts for some functional computer programs.

III. Copyright Protection for Information Stored in DNA Compounds

Scholars have speculated about the relationship between copyright and DNA for decades.[81] Previous articles addressing the relationship between DNA and

[77] It was estimated in 2015 that all of the worlds' digital information could be stored in 9 liters of DNA solution. *See*, John Markoff, *Synthetic DNA is Seen as Way to Store Data for Centuries*, N.Y. TIMES, Dec. 3, 2015, at B1.

[78] Julian Koch, Silvan Gantenbein, Kunal Masania, Wendelin J. Stark, Yaniv Erlich, Robert N. Grass, *A DNA storage architecture to create materials embedded with memory*, 38 NATURE BIOTECH. 39, 39 (2020), https://www.nature.com/articles/s41587-019-0356-z.epdf?sharing_token=c2Pos7WwuJ9MtMGj-qdFQNRgN0jAjWel9jnR3ZoTv0M4Woj1cE3OBfuw5I5lxno_c7GoY2-6n89GH-ivEpAEquXjCDjHA8AZdMxio1_l4363ezz9mt81f8Ux0-ThMicOcJ3jN17Y29Zjoyaxwr_igBrd3adSox6_-oH3cNq6DJ1ULcp4ByGA1x5klZvn7uokkqRdDbeKAWtoswVNzrPPjDH9ZvIMjCntNYKs1wFv1aqEveze5ycUK_kclvwA58FZVaUZfx68IiSptq24UNUNJR2zrvMR0Vwz_5wNAwNHnqY%3D&tracking_referrer=spectrum.ieee.org (storing data in a transparent lens and a bunny shaped figurine).

[79] Sang Yup Lee, *DNA Data Storage Is Closer Than You Think*, SCI. AM. (July 1, 2019), https://www.scientificamerican.com/article/dna-data-storage-is-closer-than-you-think/ (DNA compounds have remained stable for 500,000 years while the magnetic or optical media for digital data is subject to degradation in less than 100 years).

[80] OFF. OF THE DIR. OF NAT'L INTELL., MOLECULAR INFO. STORAGE (MIST), (last visited Feb. 9, 2021), www.iarpa.gov/index.php/research-programs/mist/mist-baa.

[81] Irving Kayton, *Copyright in Genetic Works*, 50 GEO. WASH. L. REV. 191(1982); Jorge A. Goldstein, *Copyrightability of Genetic Works*, NATURE BIOTECH. 138 (Feb. 1984); Donna Smith, Comment, *Copyright Protection for the Intellectual Property Rights to Recombinant Deoxyribonucleic Acid: A Proposal*, 19 ST. MARY'S L.J. 1083 (1988); Dan L. Burk Copyrightability of Recombinant DNA Sequences, 29 Jurimetrics J. 469 (1988-89). Some have also posed the related

copyright have treated human-constructed DNA as if it constitutes a new type of "work" which may or may not be entitled to copyright protection under the statute.[82] Some proponents of copyright protection have argued that because DNA operates within cells as computer programs operate in computers, DNA sequences, like computer programs, are works encompassed by the definition of "literary works" in the Copyright Act and therefore copyrightable.[83] Others have argued that, alternatively, if DNA sequences are not considered literary works, they are still within the scope of works that Congress intended copyright to protect.[84] The Copyright Office invokes the same premise but arrives at the opposite conclusion, finding that works such as "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."[85]

---

question of copyright protection for genetically engineered organisms. *See*, Irving Kayton, *Copyright in Genetic Works*, 50 GEO. WASH. L. REV. 191, 218 (1982).\;); For a discussion of copyright for genetically engineered organisms, *see*, Michael D. Murray, Post-Myriad Genetics Copyright of Synthetic Biology and Living Media, 10 Okla. J.L. & Tech. 1 (2014). Dan L. Burk, DNA Copyright in the Administrative State, 51 U.C. Davis Law Review (2018); Andrew Torrance, *Synthesizing Law for Synthetic Biology*, 11(2) MINN. J. L., SCI. AND TECH. 629, 642-48 (2010). Christopher M. Holman, Claes Gustafsson & Andrew W. Torrance, *Are Engineered Genetic Sequences Copyrightable?: The U.S. Copyright Office Addresses a Matter of First Impression,* 35 BIOTECH. L. REP. 103 (2016); Christopher M. Holman, Charting the Contours of a Copyright Regime Optimized for Engineered Genetic Code, 69:3 Oklahoma L. REV. 399, 402 (2017).

[82] *See* Dan L. Burk, *Copyrightability of Recombinant DNA Sequences*, 29 Jurimetrics J. 469, 495 (1989) ("[B]oth commentators favoring rDNA copyright and those opposing it concede that, with regard to the inclusion of a new category of works under the statute, intent may be determined by analogy to the enumerated categories.").

[83] Christopher M. Holman, Claes Gustafsson & Andrew W. Torrance, *Are Engineered Genetic Sequences Copyrightable?: The U.S. Copyright Office Addresses a Matter of First Impression,* 35 BIOTECH. L. REP. 103, 113 (2016).

[84] Irving Kayton, *Copyright in Genetic Works*, 50 GEO. WASH. L. REV. 191, 200-01 (1982). (asserting that even if "genetically engineered works" are not literary works, they may still be works of authorship because such works are "not 'completely outside the present congressional intent" (quoting H.R REP. NO. 1476, 94th Cong., 2d Sess. 51, *reprinted in* 1976 U.S. CODE CONG. & AD. NEWS 5669,5664)).

[85] Christopher M. Holman, Claes Gustafsson & Andrew W. Torrance, *Are Engineered Genetic Sequences Copyrightable?: The U.S. Copyright Office Addresses a Matter of First Impression,* 35

But DNA, because it functions as an information technology, does not create a new type of "work of authorship," either within the enumerated categories of the Copyright Act or outside the scope that Congress intended. Rather, DNA compounds are not works of authorship at all. DNA compounds are the physical material in which copyrightable works or other information may be fixed, in other words, "copies" under the Copyright Act.[86] It makes no more sense to ask whether DNA is copyrightable than to ask whether marks on a page are copyrightable. The *information contained in those marks* determines whether unauthorized copying is prohibited. Marks on a page organized to create a novel are clearly protected by copyright. Marks on a page that are merely tally marks used to score a game are not. The Copyright Act may grant exclusive rights to DNA compounds storing original works of authorship, but no exclusive rights should be granted to genetic DNA compounds, which store the processes of protein synthesis.

A. The Difficulty of Separating the Information from the Object

Fundamental to the current incarnation of the copyright statute is the notion that copyright protects "works."[87] According to the statute, copyright subsists in works when they are fixed in a "tangible medium of expression" (in the words of Section 102(a)) or a "material object" (in the words of Section 101).[88] Literary works are often fixed as ink on paper. Motion pictures are often fixed as patterns of light sensitive chemicals on film. Sound recordings are often fixed as patterns of magnetic particles on tape. Of course, each of these works is also often fixed in

---

BIOTECH. L. REP. 103, 122 (2016). *See also* U.S. COPYRIGHT OFF., COMPENDIUM OF U.S. COPYRIGHT OFF. PRACTICES § 313.3(A) (3d ed, 2021).

[86] 17 U.S.C. 101 ("'Copies' are material objects, other than phonorecords, in which a work is fixed by any method now known or later developed, and from which the work can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device. The term 'copies' includes the material object, other than a phonorecord, in which the work is first fixed.")

[87] 17 U.S.C. §102(a) ("copyright protection subsists . . . in original works of authorship fixed in any tangible medium of expression"). *See also* Rotstein, Robert, *Beyond Metaphor: Copyright Infringement and the Fiction of the Work*, Chicago-Kent L. R. 68:725, 738 (1993) (The "autonomous work that is the product of authorial originality [is] an idea central to the current system of copyright.").

[88] Presumably these are two terms for the same thing.

digital form as a pair of electrical states in computer software.

The Copyright Act and its legislative history make clear that a copyrightable "work" is separate and distinct from the physical form in which it exists. For example, the definition of a "literary work" explicitly disregards "the nature of the material objects . . ." in which such works are embodied.[89] Copyright protection subsists in a novel whether it is fixed in a printed book or in a digital file. As stated in the Register of Copyright's Supplementary Report written during drafting of the Copyright Act, "A consistent effort has been made in this section and throughout the bill to distinguish between the 'original work' which is the product of the author's creative intellect and which is the real subject of copyright protection and . . . the material objects embodying the work[]."[90]

Before the onset of copyright jurisprudence addressing computer programs, this concept was relatively settled. A sound recording stored on tape is not copyrightable because magnetic tape is copyrightable. The work is the music people hear. The magnetic tape is simply the tangible medium in which the work is fixed. The work is also distinct from the pattern of magnetic particles on the tape. The characteristics of the information stored in that pattern determines whether the tape may be copied without the permission of a copyright holder. Similarly, with respect to literary works, copyright does not depend on the physical form in which they exist. Not all letters or words printed on paper are entitled to copyright protection. A list of names and phone numbers on a piece of paper is not a copyrightable work of authorship.[91] A poem clearly is. A novel is copyrightable whether it is stored with ink on paper or carved into stone.

However, in the context of computer programs, the distinction between the "work" and its physical form has blurred. Courts often fail to distinguish between the potentially copyrightable information and the media in which it is fixed. Court decisions refer interchangeably to computer programs, to the computer software in which such programs are stored, and to the computer code in which such programs

---

[89] 17 U.S.C. §101.
[90] HOUSE COMM. ON THE JUDICIARY, 89TH CONG., 1ST SESS., COPYRIGHT LAW REVISION PART 6: SUPPLEMENTARY REPORT OF THE REGISTER OF COPYRIGHTS ON GENERAL REVISION OF THE U.S. COPYRIGHT LAW. at xxxii.
[91] Feist Publications, Inc., v. Rural Telephone Serv. Co., 499 U.S. 340 (1991).

are described or written.[92] For example, the Supreme Court's decision in *Google v. Oracle* at one point states that Google copied a portion of a program[93] and at another states that Google copied code.[94] The Court also appears to equate software with code and programs.[95] Lloyd Weinreb has discussed the failure of the court in *Lotus Development Corp. v. Paperback Software International* "to distinguish the program from the code. . . ."[96] Similarly, the CONTU Report, relied on by Congress to establish the scope of copyright protection for computer programs, variably equates computer programs with copyrightable works[97] and with tangible media of expression.[98]

Previous commentary on copyright and DNA similarly fails to distinguish among the physical chemical compound, the information stored in the sequence of nucleotides in the compound, and the series of letters often used to represent the sequence of nucleotides. Discussions of copyright and DNA variably focus on whether "genetic works,"[99] genetically engineered organisms,[100] "engineered

---

[92] For a discussion of computer programs (and genetic DNA compounds) as described in code or existing as code, *see* Section IV, *infra*.
[93] Google, LLC v. Oracle America, Inc., 593 U.S. _ at 1 (2021).
[94] Google, LLC v. Oracle America, Inc., 593 U.S. _ at 3 (2021).
[95] Google, LLC v. Oracle America, Inc., 593 U.S. _ at 1 (2021) (referring to software as being "written" and software as carrying out tasks).
[96] Lloyd Weinreb, *Copyright for Functional Expression*, 111 HARV. L. REV. 1149, 1155-1157 (1998).
[97] *See*, CONTU REPORT at 21-2. (When either a program or a motion picture is used in conjunction with a properly working machine, "the same result will occur on the first, the second, or the thousandth running."); CONTU REPORT at 21 (Computer programs should be treated as copyrighted written rules to a game.).
[98] *See*, CONTU REPORT at 21-2 (When either a program or a phonorecord is used in conjunction with a properly working machine, "the same result will occur on the first, the second, or the thousandth running."); CONTU REPORT at 21 ("Programs should no more be considered machine parts than videotapes should be considered parts of projectors or phonorecords parts of sound reproduction equipment.").
[99] Irving Kayton, *Copyright in Genetic Works*, 50 GEO. WASH. L. REV. 191, 201 (1982). (Engineered genetic works should be copyrightable.).
[100] *See*, Irving Kayton, *Copyright in Genetic Works*, 50 GEO. WASH. L. REV. 191, 218 (1982).\; Jorge A. Goldstein, *Copyrightability of Genetic Works*, NATURE BIOTECH. 138, 139 (Feb. 1984); Michael D. Murray, Post-Myriad Genetics Copyright of Synthetic Biology and Living Media, 10 Okla. J.L. & Tech. 1 (2014).

genetic sequences,"[101] "recombinant DNA molecules,"[102] "engineered genetic code,"[103] "human-designed DNA,"[104] "recombinant DNA sequences,"[105] "DNA code,"[106] "DNA molecules,"[107] or simply "DNA"[108] are copyrightable. Rarely, if ever, are these terms defined or the differences between them discussed. In fact, they are often used interchangeably.[109] But, the distinction is important. It is difficult to resolve the complicated question of copyright protection in the context of new technologies without a commonly understood vocabulary.[110] An article posing the question of whether "DNA molecules" are copyrightable appears to be

---

[101] Christopher M. Holman, Claes Gustafsson & Andrew W. Torrance, *Are Engineered Genetic Sequences Copyrightable?: The U.S. Copyright Office Addresses a Matter of First Impression*, 35 BIOTECH. L. REP. 103 (2016).

[102] Donna Smith, Comment, *Copyright Protection for the Intellectual Property Rights to Recombinant Deoxyribonucleic Acid: A Proposal*, 19 ST. MARY'S L.J. 1083, 1096, 1106 (1988) (Because recombinant DNA molecules could be viewed as "machine readable," they should be copyrightable just as machine readable computer programs.).

[103] Christopher M. Holman, Charting the Contours of a Copyright Regime Optimized for Engineered Genetic Code, 69:3 Oklahoma L. REV. 399, 402 (2017).

[104] Christopher M. Holman, Claes Gustafsson & Andrew W. Torrance, *Are Engineered Genetic Sequences Copyrightable?: The U.S. Copyright Office Addresses a Matter of First Impression*, 35 BIOTECH. L. REP. 103, 118 (2016) ("[T]he justification for maintaining copyright protection for software while denying it for human-designed DNA becomes increasingly questionable.")

[105] Dan L. Burk Copyrightability of Recombinant DNA Sequences, 29 Jurimetrics J. 469 (1988-89).

[106] Devdatta Malshe, *Copyrighting DNA: An Off-Label Use*, 19 WAKE FOREST J. OF BUS. AND INTELL. PROP. L. 34, 42 (2018).

[107] Andrew W. Torrance, *DNA Copyright*, 46 VAL. U. L. REV. 1, 35 (2011) ("DNA molecules are copyrightable. . ..").

[108] Andrew W. Torrance, *DNA Copyright*, 46 VAL. U. L. REV. 1 (2011) (title).

[109] *See*, Christopher M. Holman, Claes Gustafsson & Andrew W. Torrance, *Are Engineered Genetic Sequences Copyrightable?: The U.S. Copyright Office Addresses a Matter of First Impression*, 35 BIOTECH. L. REP. 103,101 (2016) (referring to the copyrightability of "engineered DNA sequences," "genetic code" and "engineered genetic sequences" interchangeably); Andrew W. Torrance, *DNA Copyright*, 46 VAL. U. L. REV. 1, 4 n. 10 (2011). ("In this article, 'gene' and 'DNA sequence' are often used interchangeably, where appropriate.").

[110] *See*, John A. Kidwell, *Software and Semiconductors: Why Are We Confused*, 70 MINN. L. REV. 533, 538-40 (1985) (analogizing the difficulties that arise when people lack a shared vocabulary when discussing computer software to the difficulties that might arise when contract law must contend with the repair of a new type of watch. "If a part in a new timepiece seems to have no analog to any part in an old watch, some watchmakers may begin to call one part a blodget, and the other a wedge, while others call the same parts widgets and wedges. . .. The question here is whether watches are like computer programs, or judges like watchmakers.").

asking whether a class of physical chemical compounds is protected by copyright. An article arguing that "DNA sequences" are copyrightable as literary works may be proposing that the series of As, Ts, Cs, and Gs that is often used to describe a DNA compound is copyrightable but not the compound itself. An article discussing "DNA code"[111] may simply be arguing that a coded clue written in As, Ts, Cs, and Gs is copyrightable. Each of these questions presents a different set of issues and a different answer. Assessing copyright protection for a work in a DNA compound or molecule is different from assessing copyright protection for a literary work representing the compound.

The Copyright Office similarly seems to have lost sight of the fundamental distinction between the information constituting the work of authorship and the physical media in which it is fixed. For example, the Copyright Office declared in its Compendium that "works" such as "DNA sequences and other genetic, biological, or chemical substances or compounds" as a rule do not constitute copyrightable subject matter.[112] It is unclear what the Copyright Office means by

---

[111] It is unclear whether DNA code refers to verbal representations of DNA compounds in the sense of computer code or whether it refers to the correspondence between the sets of three nucleotides and the amino acids that make up proteins. *See* Assoc. for Molecular Pathology v. U.S. Pat. and Trademark Off., 702 F. Supp. 2d 181, 193-200 (S.D.N.Y. 2010), aff'd in part & rev'd in part, 689 F.3d 1303 (Fed. Cir. 2012), aff'd in part & rev'd in part, 569 U.S. 576 (2013).

[112] U.S. COPYRIGHT OFF., COMPENDIUM OF U.S. COPYRIGHT OFF. PRACTICES § 313.3(A) (3d ed, 2021). The Office arrived at this conclusion despite the fact that Congress has commissioned no series of extensive studies as it did with respect to digitally stored information. Indeed, the Compendium provides no support for such a blanket statement other than conclusively stating that such sequences are examples of the ideas, procedures, processes, systems, methods of operation, concepts, principles, or discoveries that are excluded from copyright protection under section 102(b) of the Copyright Act. A letter from Robert J. Kasunic, Associate Register of Copyrights and Director of Copyright Policy and Practices in response to a request by Christopher Holman and Andrew Torrance and Dr. Claes Gustafsson for reconsideration of a refusal to register a specific human-constructed DNA compound labeled the "Prancer DNA Sequence" provides a window into the reasoning behind the statement in the Compendium. The letter states that 1) the Prancer DNA Sequence is not "within the congressionally established categories of authorship in title 17;" 2) the Prancer DNA Sequence does "not include a sufficient quantum of copyrightable authorship," and 3) copyright protection is precluded for the Prancer DNA Sequence because protection does not extend to any idea, procedure, process, system, method of operation, principle, or discovery. February 11, 2014 letter from Robert J. Kasunic to Howard Simon reprinted as Supplementary Document 2: Affirmance of Refusal for Registration in Christopher M. Holman, Claes Gustafsson & Andrew W.

the term "DNA sequences," but "genetic, biological, or chemical substances or compounds" such as DNA compounds are not a class of works of authorship that potentially constitute copyrightable subject matter. They are not works at all. "Substances and compounds" are tangible media. As explained above, they are "copies" in which a copyrightable work or some other creation may be fixed. To determine whether copyright prohibits the copying of a particular DNA compound, one should examine the information, the potential "work of authorship," stored in it.

The difficulty in conceiving of the information stored in a DNA compound separately from the tangible compound itself is not limited to the context of copyright. In the seminal patent case *Association for Molecular Pathology v. Myriad Genetics*, one disagreement between the parties centered around whether the term "sequence" in Myriad's patent claims referred to, as Myriad wrote in its brief on appeal, "mere information" rather than a physical molecule.[113] The issue was important because, as the Supreme Court recognized, the value of a patent claim in DNA is attributable more to the information contained in the sequence of nucleotides in the DNA than to "the specific chemical composition of a particular molecule."[114] Unfortunately, the Supreme Court's holding in the case did little to settle that particular dispute.[115] The Court's decision states that, despite decades in which the Patent Office granted patents for DNA sequences consisting of partial genes, "genes *and* the information they encode are not patent eligible . . . ."[116]

More than ten years earlier, comments by Rebecca Eisenberg at a Boston University symposium panel entitled *Molecules vs. Information: Should Patents*

Torrance, *Are Engineered Genetic Sequences Copyrightable?: The U.S. Copyright Office Addresses a Matter of First Impression,* 35 BIOTECH. L. REP. 103 (2016).

[113] Brief for the Appellants at 58, Assoc. for Molecular Pathology v. Myriad Genetics, Inc., 569 U.S. 576, 596 (2013).

[114] Assoc. for Molecular Pathology v. Myriad Genetics, Inc., 569 U.S. 576 (2013).

[115] In contrast to the lack of precision of United States Supreme Court's reasoning in *Myriad,* the Federal Court of Australia distinguished between the genetic information stored in DNA and DNA as a tangible material. The Australian court stated that patent claims to the tangible DNA compound "could never be infringed by someone who merely reproduced a DNA sequence in written or digitized form." Cancer Voices Australia v. Myriad Genetics, Inc. 99 IPR 576 (2013) at para. 76.

[116] Assoc. for Molecular Pathology v. Myriad Genetics, Inc., 569 U.S. 576, 596 (2013) (emphasis added).

Protect Both? foreshadowed the resulting reversal of fortunes for biotechnology companies. According to Eisenberg, squeamishness about patenting DNA sequences was largely due to a feeling that patenting them is akin to patenting information. As part of a panel discussing whether patents should protect both DNA molecules and the information they contain, Eisenberg recognized, "in the early days, patenting genes looked like patenting drugs," but since the attempts of the National Institutes of Health to patent expressed sequence tags[117] during the Human Genome Project, "it looks more like patenting scientific information."[118]

But no matter how difficult the task, in order to determine the proper scope of copyright protection, the underlying information must be identified separately from the physical medium in which it is fixed in order to evaluate the potentially copyrightable "original work of authorship." As explained by the House Report to the Copyright Act, the nature of the physical form in which the information is fixed does not determine the copyrightability of the work.[119] The failure to separate the DNA compound, itself, from the information stored in DNA may be a roadblock to recognizing that DNA is simply a new type of material object in which a copyrighted work may be fixed rather than a new type of copyrightable work. In addition, failing to distinguish between the physical molecule and the information that is the copyrightable work can lead to erroneous conclusions. If the work protected by copyright is confused with the compound in which that information is stored, a finding that copyright prohibits the reproduction of one DNA compound could lead to the conclusion that copyright prohibits the reproduction of DNA compounds in general.[120] If the literary work describing or representing a DNA compound is confused with the compound itself, copyright protection for a series of As, Ts, Cs, and Gs could lead to the extension of copyright protection to the DNA compound described by that string of letters.[121]

---

[117] Expressed sequence tags, or ESTs are polynucleotide molecules that have the nucleotide sequence of a short segment of cDNA.
[118] Rebecca Eisenberg, *Molecules vs. Information: Should Patents Protect Both?,* 8 B.U. J. SCI. & TECH. L. 190, 191 (2002).
[119] H.R. REP. NO. 94-1476, at 53 (1976), *re- printed in* 1976 U.S.C.C.A.N 5659, 5666.
[120] *See* Section III.C., *infra*.
[121] *See* Section IV, *infra*.

B. Copyright Subsists in Original Works of Authorship Fixed in DNA Compounds

Although the Copyright Act clearly states that copyright protection subsists in "original works of authorship, the statute lacks a definition of those "works," the very creation in which copyright subsists. The failure to define the work was not an oversight. Indeed, the legislative history to the Act states that "the phrase 'original works of authorship,' [] is purposely left undefined . . . ."[122] The only guidance provided by the statute is a list of eight "illustrative but not limitative"[123] categories of "works of authorship": literary works; musical works; dramatic works; pantomimes and choreographic works; pictorial, graphic and sculptural works; motion pictures and other audiovisual works; sound recordings; and architectural works.[124]

As described in Section II, above, a sonnet and a manuscript,[125] sculpted

[122] H.R. REP. NO. 1476, 94th Cong., 2d Sess. 47, at 51 (1976). S. REP. NO. _ at 50.
[123] H.R. REP. NO. 1476, 94th Cong., 2d Sess. 47 (1976). S. REP. NO. _. *See also* 1-2 Nimmer on Copyright §2.03.
[124] 17 U.S.C. § 102.
[125] Nick Goldman, *et al.*, *Towards practical, high-capacity, low-maintenance information storage in synthesized DNA*, 494 NATURE 77, 77 (2013). Luis Ceze, et al., *Molecular digital data storage using DNA*, 20 NATURE REV. GENETICS 456, 459 fig. 3 (2019); Susan Young Rojahn, *An Entire Book Written in DNA*, MIT TECHNOLOGY REVIEW (Aug. 16, 2012), https://www.technologyreview.com/2012/08/16/184447/an-entire-book-written-in-dna/.

bunny figurines,[126] a recording of a musical performance,[127] photographs,[128] and movies[129] have each been stored in human-constructed DNA compounds. If original, such literary works, sculptural works, sound recordings, graphical works and motion pictures surely qualify as protectable works of authorship under the Copyright Act. Counter to the guidance provided by the Copyright Office's Compendium,[130] which would disqualify works of authorship from copyright protection simply because they are fixed as a DNA compound, the copyright protection afforded to any one of these creations should not, and practically cannot, depend on the tangible media of expression in which they are fixed. The fact that a work is embodied in a DNA compound does not exclude it from copyright protection any more than the fact that a work is embodied in ink on a page should compel the result that it is entitled to copyright protection.

A DNA compound is simply a new tangible medium of expression in which works may be permanently and stably fixed. According to the statute, copyright subsists in works when they are fixed in a "tangible medium of expression" as required by the statute.[131] The statute explains that a work is "fixed" in a material

[126] A DNA storage architecture to create materials embedded with memory, 38 Nature Biotech. 39-43 (2020). https://www.nature.com/articles/s41587-019-0356-z.epdf?sharing_token=c2Pos7WwuJ9MtMGj-qdFQNRgN0jAjWel9jnR3ZoTv0M4Woj1cE3OBfuw5I5lxno_c7GoY2-6n89GH-ivEpAEquXjCDjHA8AZdMxio1_l4363ezz9mt81f8Ux0-ThMicOcJ3jN17Y29Zjoyaxwr_igBrd3adSox6_-oH3cNq6DJ1ULcp4ByGA1x5klZvn7uokkqRdDbeKAWtoswVNzrPPjDH9ZvIMjCntNYKs1wFv1aqEveze5ycUK_kclvwA58FZVaUZfx68IiSptq24UNUNJR2zrvMR0Vwz_5wNAwNHnqY%3D&tracking_referrer=spectrum.ieee.org

[127] *#MemoriesInDNA portrait project blends DNA technology and art to memorialize pioneering scientist Rosalind Franklin*, ALLEN SCHOOL NEWS, U. OF WASH. (Feb. 24, 2020), https://news.cs.washington.edu/2020/02/24/memoriesindna-portrait-project-blends-dna-technology-and-art-to-memorialize-pioneering-scientist-rosalind-franklin/

[128] Nick Goldman, *et al.*, *Towards practical, high-capacity, low-maintenance information storage in synthesized DNA*, 494 NATURE 77, 77 (2013).

[129] Ed Yong, *Scientist Can Use CRISPR to Store Images and Movies in Bacteria*, THE ATL. (July 12, 2017), https://www.theatlantic.com/science/archive/2017/07/scientists-can-use-crispr-to-store-images-and-movies-in-bacteria/533400/.

[130] U.S. COPYRIGHT OFF., COMPENDIUM OF U.S. COPYRIGHT OFF. PRACTICES § 313.3(A) (3d ed, 2021).

[131] 17 U.S.C. § 102(a).

object[132] when it "is sufficiently permanent or stable to permit it to be perceived, reproduced or otherwise communicated for a period of more than transitory duration."[133] Fixing a work of authorship as nucleotides in a DNA compound allows it to be perceived and reproduced as required by the statute. With the aid of DNA sequencing machines and computers, the order of nucleotides in a DNA compound may be "read" and translated first into digital form and then into a manuscript, musical performance or movie that can be perceived by humans. DNA compounds can be reproduced either chemically or within cells.[134] DNA compounds are extremely stable, more stable than any available digital technology.[135] Scientists, with the aid of DNA sequencing technology, have been able to sequence or "read" the sequence of nucleotides in the DNA from a 700,000-year-old horse bone fragment.[136] The legislative history of the Copyright Act clearly shows that the drafters contemplated that a work may be fixed in a medium that did not exist at the time of the drafting.[137] In addition, the statute explicitly states that perception and reproduction with the aid of a machine or device is sufficient.[138]

Under the Copyright Act the owner of the copyright in a sonnet should have the exclusive right to reproduce the work in DNA as well as in paper copies.[139] The same should be true of any other copyrightable work fixed in a DNA compound.

---

[132] Such material objects are defined as "copies" under the Copyright Act. 17 U.S.C. §101 (defining "copies" as "material objects in which a work is fixed by any method now known or later developed from which a work can be perceived or reproduced, either directly or indirectly, with the aid of a machine or device.").

[133] Definition of "fixed" in 17 U.S.C. § 101.

[134] Luis Ceze, *et al.*, *Molecular digital data storage using DNA*, 20 NATURE REV. GENETICS 456, 459 fig. 3 (2019).

[135] Luis Ceze, et al., *Molecular digital data storage using DNA*, 20 NATURE REV. GENETICS 456, 456 (2019).

[136] Craig D. Millar & David M. Lambert, *Towards a million-year-old genome*, 499 NATURE 34 (2013). https://doi.org/10.1038/nature12263

[137] *See*, 17 U.S.C. §101 (defining copies as "material objects in which a work is fixed by any method now known or later developed. . ..").

[138] *See*, 17 U.S.C. §101 (defining copies as material objects "from which a work can be perceived or reproduced, either directly or indirectly, with the aid of a machine or device."). *See also* H.R. Rep. No. 94-1476, at 53 (1976), *re- printed in* 1976 U.S.C.C.A.N 5659, 5666 (A copy is intended "to comprise all of the material objects in which copyrightable works are capable of being fixed.").

[139] 17 U.S.C. §106.

Copyright protection for works such as photographs, sculptures, motion pictures or novels stored in DNA compounds is equivalent to copyright protection for such works in digital information technology. Copyright should subsist in a motion picture whether it is fixed on film, video tape, DNA, or a digital storage device such as a DVD.

However, it is important to keep in mind that this conclusion implies nothing about the "copyrightability of DNA" because DNA is not the work. The information stored in the form of the sequence of nucleotides in a DNA compound are the works: the literary work, the motion picture, and the sound recording. The DNA compound is simply the tangible medium in which they are fixed. If the information stored in a DNA compound is kept conceptually separate from the physical DNA compound, it is clear that copyright subsists in an original literary work, motion picture, or sound recording fixed as a series of nucleotides in a DNA compound. The original work of authorship is the literary work, motion picture or sound recording. The work is fixed in the tangible medium of DNA. According to the definitions contained in the statute, the DNA compound in which that work is fixed is a "copy"[140] of that work. If the information stored in a DNA compound is an original literary work, motion picture or sound recording, Section 106 of the Copyright Act grants the exclusive right to reproduce and distribute that work in DNA compounds.

C. Copyright and Functional Information Stored in DNA Compounds

Although a DNA compound may be a "copy" in which a copyrightable work is embodied, that fact does not compel the conclusion that anything embodied in a DNA compound is a copyrightable work.[141] As recognized in the House Report accompanying the bill which led to the 1976 Copyright Act, "[i]t is possible to have an 'original work of authorship' without having a 'copy' or 'phonorecord' embodying it, and it is also possible to have a 'copy' or 'phonorecord' embodying

[140] 17 U.S.C. §101.
[141] Indeed, to do so would be to revert to earlier copyright statutes in which works that take certain forms are entitled to copyright protection. *See* Copyright Act of May 31, 1790, ch. 15, § 1, 1 Stat 124, 124 (repealed 1831) (granting copyright protection to books, maps, and charts).

something that does not qualify as an 'original work of authorship.'"[142] How then should we analyze copyright protection for genetic DNA, which embodies the process of protein synthesis rather than a novel or motion picture?

Prior scholarship has premised the possibility of copyright protection for genetic DNA on an analogy between DNA and digital technology.[143] For example, genes have been equated with computer software,[144] "engineered genetic works" have been compared to computer programs,[145] and DNA compounds have been viewed as computer programs fixed in software.[146] The logic of copyright

[142] H.R. REP. NO. 94-1476, at 53 (1976), *re- printed in* 1976 U.S.C.C.A.N 5659, 5666.

[143] Irving Kayton, *Copyright in Genetic Works*, 50 GEO. WASH. L. REV. 191, 201 (1982) (Engineered genetic works "are certainly analogous, if not nearly identical, to computer programs . . . [and b]ecause of this similarity . . . should be copyrightable."); Christopher M. Holman, Claes Gustafsson & Andrew W. Torrance, *Are Engineered Genetic Sequences Copyrightable?: The U.S. Copyright Office Addresses a Matter of First Impression*, 35 BIOTECH. L. REP. 103, 118 (2016) ("[T]he justification for maintaining copyright protection for software while denying it for human-designed DNA becomes increasingly questionable."); Andrew Torrance, Synthesizing Law for Synthetic Biology, 11 MINN. J. L., SCI. AND TECH. 629, 648 (2010) ("[S]ynthetic 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."); Dan L. Burk, *Copyrightability of Recombinant DNA Sequences*, 29 JURIMETRICS J. 469, 472 (1989) ("[A]dvocates on both sides of the DNA copyright debate have discussed the analogy between computer software and recombinant DNA. . .."); Jorge A. Goldstein, *Copyrightability of Genetic Works*, NATURE BIOTECH. 138, 139 (Feb. 1984) ("The strongest reason for arguing . . . that polynucleotide molecules [including DNA] are appropriate media of expression for genetic works is by analogy with the computer world."); Donna Smith, Comment, *Copyright Protection for the Intellectual Property Rights to Recombinant Deoxyribonucleic Acid: A Proposal*, 19 ST. MARY'S L.J. 1083, 1096, 1106 (1988) (Describing "[t]he issues surrounding the scope of protection of a copyrighted computer program" as "also pertinent to copyrighted DNA.");

[144] Andrew Torrance, Synthesizing Law for Synthetic Biology, 11 MINN. J. L., SCI. AND TECH. 629, 648 (2010) ("[S]ynthetic 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.").

[145] Irving Kayton, *Copyright in Genetic Works*, 50 GEO. WASH. L. REV. 191, 201 (1982). (Engineered genetic works "are certainly analogous, if not nearly identical, to computer programs . . . [and b]ecause of this similarity . . . should be copyrightable.").

[146] Donna Smith, Comment, Copyright Protection for the Intellectual Property Rights to Recombinant Deoxyribonucleic Acid: A Proposal, 19 St. Mary's L.J. 1083, 1096, 1106 (1988) (Because recombinant DNA molecules could be viewed as "machine readable," they should be copyrightable just as machine readable computer programs.).

protection for computer code has been applied to "DNA code." [147] But arguing for copyright protection for engineered genetic works or genes based on an analogy to computer programs risks importing the same logical missteps that plague copyright protection for computer programs.

One misconception that plagues analyses of copyright protection for computer programs is the mischaracterization of computer programs as "instructions." Even the Copyright Act itself defines a computer program as "statements or instructions to be used directly or indirectly in a computer in order to bring about a certain result.".[148] But as recognized by Lloyd Weinreb in discussing computer programs, "there is nothing that can be described as 'statements or instructions' except as an elaborate metaphor."[149] Rather, a computer program is more accurately compared to a part of a machine or the process carried out by a machine. Therefore, as Weinreb notes, "a program is not copyrightable, any more than are the gears that operate the shift of a car or the shifting of the gears itself . . . ."[150]

DNA sequences have been similarly characterized as "instructions."[151] But

---

[147] Devdatta Malshe, *Copyrighting DNA: An Off-Label Use*, 19 WAKE FOREST J. OF BUS. AND INTELL. PROP. L. 34, 42 (2018) (questioning the logic of protection for "computer code" while denying it to "DNA code").

[148] 17 U.S.C. § 101. Lloyd Weinreb points out that the definition of computer program in the Copyright Act as "instructions to be used directly or indirectly in a computer in order to bring about a certain result" at the very least confuses the computer code with the computer program. "Although the description of programs as 'statements or instructions' plainly refers to the program code, the reference to use 'in a computer' can refer only to the program in operation." Lloyd Weinreb, *Copyright for Functional Expression*, 111 HARV. L. REV. 1149, 1157 (1998).

[149] Lloyd Weinreb, *Copyright for Functional Expression*, 111 HARV. L. REV. 1149, 1157 (1998).

[150] Lloyd Weinreb, Copyright for Functional Expression, 111 Harv. L. Rev. 1149, 116**8** (1998). In this quotation, Professor Weinreb appears to be treating the "program" either as fixed in hardware or software (in which case it is analogous to the gears) or as the process operating in the computer (in which case it is analogous to the shifting of the gears). The statement is true in either case.

[151] *See*, IVER P. COOPER, BIOTECH. AND THE LAW, § 14.6 (May 20, 2015) ("A DNA nucleotide sequence may be compared to a rule book (a set of instructions for playing a game), to a recipe (a set of instructions for making a complex chemical substance), and to a blueprint or architectural plan (a set of instructions for constructing a physical structure)."); Andrew W. Torrance, *DNA Copyright*, 46 VAL. U. L. REV. 1, 33 (2011) (""A gene is a set of instructions for producing a polypeptide.").

just as computer programs are not instructions, the information in DNA compounds does not function as instructions for the processes carried out in cells. There is no one to instruct.[152] The sequence of nucleotides in a DNA compound does not describe the technology or provide instructions on how to use it. Rather, just as a "computer carries out the program by means of the flow of current through electronic circuitry [and] needs no instructions (and could follow none were they given),"[153] a cell carries out the process of protein synthesis by means of chemical reactions and neither needs nor is capable of following instructions.

To apply Weinreb's language to the context of DNA, DNA "requires us to replace our concrete conception of a machine as a physical object with the abstraction of a means to perform a function."[154] In the context of patent protection for DNA technology, both courts and commentators have been more adept at recognizing the role of DNA compounds in that abstraction. As the Federal Circuit has recognized, rather than existing as passive instructions followed by actors in a process, "[t]he majority of genes *act* by guiding the production of polypeptide chains that form proteins."[155] In discussing the application of patent law to modern biotechnology, Dan Burk has suggested that "[r]ather than comparisons to blueprints and [instructions], DNA might better be compared to a cog in a [genetic information expression] machine. . . ."[156]

If genetic DNA compounds are cogs in cellular machinery, how should we

[152] *See*, Pamela Samuelson, *Contu Revisited: The Case Against Copyright Protection for Computer Programs in Machine-Readable Form*, 1984 DUKE L. J. 663, 727 (1984). (distinguishing between a book which contains a set of instructions and a computer program in machine-readable form which contains a set of instructions by clarifying that "[t]he former informs a human being about how the task might be done; the latter does the task.")
[153] Lloyd Weinreb, *Copyright for Functional Expression*, 111 HARV. L. REV. 1149, 1157 (1998).
[154] Lloyd Weinreb, *Copyright for Functional Expression*, 111 HARV. L. REV. 1149, 1169 (1998).
[155] Assoc. for Molecular Pathology v. U. S. Pat. and Trademark Off., 689 F.3d 1301, 1310 (2012), aff'd in part & rev'd in part, 133 S. Ct. 2107 (2013) (emphasis added).
[156] Dan L. Burk, *The Problem of Process in Biotechnology*, 43 HOUS. L. REV. 561, 583 (2006) (As Burk points out, this genetic information expression machine in which DNA is a cog is analogous to "Babbage's famous 'difference engine,' the conceptual precursor to modern computing, which was intended to accomplish complicated numerical calculations by means of mechanical gears.")

think about the information stored in those compounds?[157] Genetic DNA still has the qualities of an information technology, but unlike the DNA compounds discussed in Section III. B., *infra*, the stored information is not a novel, sculpture or sound recording. Rather, genetic DNA compounds contain information that can be more accurately described either as 1) an alternative representation of a protein or 2) the process of protein synthesis (the shifting of the gears).[158] This dual nature of the stored information is not unique to DNA. Indeed, any information that is perceptible only through the aid of a machine can be considered either a representation of the output of the machine or part of a process conducted by the machine which produces the output. The information stored on a magnetic tape of a sound recording may be considered either a representation of the sound recording or part of the sound producing process conducted by a tape recorder.

More specifically, the information stored in a genetic DNA compound can be considered a representation of the protein (or proteins) that would result from the chemical processes of transcription and translation that occur in a cell. The "genetic code," the direct correspondence of the sequence of nucleotides in a DNA compound to the sequence of amino acids in a protein, allows the nucleotide sequence to provide all the information necessary to construct the protein. As recognized by Eisenberg, in naturally occurring living organisms, "[o]ne can think of DNA as a tangible storage medium for information . . . about the structure of proteins."[159] In other words, for the purpose of assessing copyright protection, the

---

[157] Burk grapples with the distinction between the physical DNA compound and the information it stores discusses this metaphysical question in attempting to differentiate between patents that claim exclusive ownership of a product versus those that claim ownership of a process. He notes that "[t]he novelty and value of biotechnological inventions [including those related to DNA] lie in their processes, which are determined by their structures. . .. But molecular structure defines the parameters for such a process, and structure falls formally into the category of products." Dan L. Burk, *The Problem of Process in Biotechnology*, 43 HOUS. L. REV. 561, 587 (2006).

[158] Following Weinreb's comparison of computer programs to the gears that operate the shift of a car, it may be tempting to alternatively consider the information stored in genetic DNA compounds as a machine part. I think that it is more accurate to consider the DNA compound or in the case of computer technology, the computer software, as being a machine part.

[159] Rebecca Eisenberg, *Molecules vs. Information: Should Patents Protect Both?*, 8 B.U. J. SCI. & TECH. L. 190, 196 (2002); Andrew Torrance recognizes the equivalence between the information stored in a genetic DNA compound and a protein when he explicitly uses DNA as shorthand for

information fixed in a DNA compound may be the protein itself.[160] It follows that because a protein is a useful object, it should be treated as any other useful object when assessing copyright protection. Unless a protein is designed for its beauty[161] as well as the function of its structure, there would be no features eligible for copyright protection as separately identifiable from the utilitarian aspects of the compound.[162] The DNA compound may be entitled to patent protection, but not copyright.

Alternatively, just as a computer program can be compared to a process, "the shifting of the gears" of a car, the information stored in genetic DNA can be considered to be part of a process: the process of protein synthesis in a cell. When an input in the form of an enzyme is introduced, it interacts with the nucleotide sequence of a genetic DNA compound causing the cell to start a series of chemical reactions to produce an output in the form of protein. The order of nucleotides in genetic DNA compounds contains the information necessary to construct the protein, complete with the information for so-called start and stop codons to start and stop the process.

Genetic DNA compounds that occur in nature, which are used by cells to produce a different protein depending on the enzyme used as an input, present a particularly compelling case for treating the information in such compounds as a process rather than an alternative form of representing a protein. Every cell in a living organism contains the same set of DNA compounds. Those compounds contain the information necessary to produce all the proteins necessary for survival of the organism. An enzyme acts as the input to the cellular "computer" which initiates the process of producing a protein. Just as the information in a computer's operating system operates to produce a different output depending on the input, the

---

DNA, RNA, and proteins. *See,* Andrew W. Torrance, *DNA Copyright,* 46 VAL. U. L. REV. 1, 26 (2011).

160 Anita Varma and David Abraham, *DNA is Different: Legal Obviousness and the Balance Between Biotech Inventors and the Market*, 9 HARV. J. OF L. & TECH. 53, 69 (1996) ("[T]he relationship between the DNA and the protein(s) it codes for, rather than the actual DNA sequence, creates value.").

161 BEAUTIFULCHEMISTRY.NET (last visited Feb. 10, 2021), https://www.beautifulchemistry.net/protein-structures.

162 *See*, 17 U.S.C. § 101.

information in a cell's DNA operates to produce a different protein depending on the enzyme introduced.[163] Because of the multiple outputs which can be produced by the process stored in such DNA compounds, the information they store seems more accurately described as the process itself rather than a representation of any one of the many possible outputs. If characterized as a process, the information stored in a genetic DNA compound is excluded from copyright protection under Section 102(b) of the Copyright Act, which explicitly prohibits the extension of copyright protection to any "procedure, process, system, or method of operation."[164]

Perhaps fortunately, distinguishing between the characterizations of the information contained in genetic DNA compounds as either an alternative representation of a protein or the process of producing a protein is not necessary to assess copyright protection. Information representing a protein should be excluded from copyright protection as utilitarian. Alternatively, if the information stored in genetic DNA compounds is characterized as the process of protein synthesis[165] then under section 102(b), copyright protection for that information is excluded and reproducing it in DNA compound "copies" is not prohibited by the Copyright Act. Thus, whether the information stored in genetic DNA is characterized as a representation of a protein or as the process by which proteins are synthesized, because of its functionality, it is excluded from copyright protection.

IV. Copyright Protection for Verbal Representations of DNA Compounds

As described in Section II, *infra*, a DNA compound is often represented as

[163] To further the analogy, DNA is stored in the nucleus or the ROM of the cellular computer. *See*, Jorge A. Goldstein, *Copyrightability of Genetic Works*, NATURE BIOTECH. 138, 140 (Feb. 1984) ("[i]t seems that DNA molecules and ROMs are not that different when viewed as functional information storage and processing media.") An enzyme works as an input which causes the cell to transfer the data stored in the nucleus (ROM) in the form of DNA to ribosomes (RAM) through the use of mRNA.

[164] 17 U.S.C. § 102(b).

[165] It is true that scientists believe that there is some portion of naturally occurring genetic DNA that has no function in the process of protein synthesis. If human-constructed genetic DNA compounds contain such nonfunctional portions, that (presumably worthless) aspect of the DNA compound could arguably be protected by copyright as expressive and not excluded by section 102(b).

a series of As, Ts, Cs, and Gs. These verbal representations of DNA compounds may be recorded on paper or in a digital database. For example, a DNA "sequence listing" required by the United States Patent Office is a long string of As, Ts, Cs, and Gs, submitted as a digital text file.[166] Verbal representations of DNA compounds also regularly appear in scientific periodicals.[167] Pharmaceutical and biotechnology companies, as well as the United States government, maintain vast databases of these verbal representations of DNA compounds. For example, the GenBank® database of the National Institutes of Health provides a searchable annotated collection of verbal representations of all publicly available DNA compounds.[168] Compilations of verbal representations of DNA compounds are used to diagnose disease,[169] place someone at a crime scene,[170] or identify someone's ethnic heritage.[171] Copyright protection for these databases as works in their entirety has already been discussed by others,[172] but in this section, I will

---

[166] *See*, MANUAL OF PATENT EXAMINING PROCEDURE §2422.03.

[167] Consider the case of a verbal representation of a DNA compound written as AATCGC and included within the context of a longer piece of written text, such as a research article published in a scientific journal. Certainly, the article is a "literary work" as defined by the copyright statute. As long as the article complies with the originality and fixation requirements, copyright subsists in the research article including the verbal representation of a DNA compound just as it does in any other scientific article. It is certainly conceivable that copying the sequence AATCGC from a scientific article disclosing the synthesis, function, and characteristics of a human constructed DNA compound with that sequence would be infringing. For example, the Southern District of New York has stated that if a defendant copied only the word SUPERCALIFRAGILISTICEXPIALIDOCIOUS in the context of the lyrics to a song "they conceivably might still be liable for infringement." Life Music, Inc. v. Wonderland Music Co., 241 F. Supp. 653, 656 (S.D.N.Y. 1965). Thus, someone copying the letters AATCGC representing a DNA compound may be liable for infringement of the copyright in the research article. As with any copyrighted work, in order to assess infringement, considerations such as the amount of the work copied, and the importance of the copied portion would need to be addressed.[167] 4 Nimmer on Copyright §13.03 (2020). In addition, the limitations on copyright protection due to functional aspects discussed in this Article should also apply.

[168] *GenBank Overview*, NCBI (last visited Feb. 10, 2021), https://www.ncbi.nlm.nih.gov/genbank/.

[169]

[170] *See generally*, Brandon L. Garrett and Peter J. Neufeld, *Invalid Forensic Science Testimony And Wrongful Convictions*, 95 Va. L. Rev. 1 (2009).

[171] Ancestry, https://www.ancestry.com/dna/ (last visited Feb. 10, 2021).

[172] *See* Ray K. Harris and Susan Stone Rosenfield, *Copyright Protection for Genetic Databases*, 45 Jurimetrics 225 (2005); M. Scott McBride, *Bioinformatics and Intellectual Property Protection*, 17 Berkeley Tech. L. J. 1331 (2002). *See also* Pamela Samuelson, *Functional*

address copyright protection for the individual verbal representations of DNA compounds.

Because the information contained in a DNA compound can more easily be searched and manipulated when it is in the form of As, Ts, Cs, and Gs than in the chemical compound, the sequence of As, Ts, Cs, Gs representing a DNA compound may be more valuable than the chemical compound itself to those who seek exclusive rights to the information.[173] This is even more true now than it was in 2001 when Rebecca Eisenberg presciently recognized that "the informational value of [DNA] sequences – by which I mean the value of simply knowing what the sequence is – is becoming more significant relative to the material value of having access to a molecule that embodies that information."[174] Now more than ever, "[k]nowing the DNA sequence . . . gives you an information base that facilitates future discoveries; and that is often . . . more significant than the tangible value of having access to the gene."[175]

Accordingly, those seeking exclusive rights to the information contained in DNA compounds have tried multiple strategies other than seeking exclusive rights to the DNA compound itself. Companies have established databases of the

---

*Compilations,*54 Houston L. R. 321 (2016). According to McBride, a scientist seeking to copy the valuable aspect of a database, the verbal representations of individual DNA compounds, "would not infringe the scientist's copyright so long as the competitor does not use the same selection or arrangement as the scientist's copyrighted database." M. Scott McBride, *Bioinformatics and Intellectual Property Protection*, 17 Berkeley Tech. L. J. 1331, 1349 (2002). (It is unclear whether McBride would arrive at the same conclusion in the context of databases of human created as well as naturally occurring DNA compounds.) Thus, any copyright protection for the individual verbal representations of DNA compounds recorded in a database would be dependent on their existence as separate and independent copyrightable works in themselves. If the DNA compounds represented in the database are human constructed, the database may instead be regarded as a collective work under the Copyright Act. *See* 17 U.S.C.§101 ("a work, . . .in which a number of contributions, constituting separate and independent works in themselves, are assembled into a collective whole.").

[173] *See generally*, Rebecca Eisenberg, *Molecules vs. Information: Should Patents Protect Both?,* 8 B.U. J. SCI. & TECH. L. 190, 196-7 (2002).

[174] Rebecca Eisenberg, *Molecules vs. Information: Should Patents Protect Both?,* 8 B.U. J. SCI . & TECH. L. 190, 196-7 (2002).

[175] Rebecca Eisenberg, *Molecules vs. Information: Should Patents Protect Both?,* 8 B.U. J. SCI . & TECH. L. 190, 198 (2002).

nucleotide sequences of DNA compounds and restricted access by offering licenses to the databases under the terms of subscription agreements.[176] Others have attempted to patent the information contained in an organism by disclosing the complete nucleotide sequence as a series of As, Ts, Cs and Gs and claiming exclusive rights to a computer medium recording the entire sequence, a fragment of the sequence or to a sequence that is at least 99.9 percent identical to the sequence.[177] Some have even claimed copyright in the verbal representation of DNA compounds.[178]

A. Verbal Representations of DNA Compounds as Literary Works

Nearly any series of letters or numbers qualifies as a "literary work" under the definition in the Copyright Act.[179] A nonsensical book of letters to be used to

---

[176] *See,* Rebecca Eisenberg, *Molecules vs. Information: Should Patents Protect Both?,* 8 B.U. J. SCI. & TECH. L. 190, 199 (2002).
[177] *Id.*
[178] *See*, Andrew W. Torrance, *DNA Copyright*, 46 VAL. U. L. REV. 1, 5 (2011) (Quoting a letter sent to a customer of a gene sequencing company claiming rights in short segments of DNA called oligonucleotides, "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.") *See also,* Stemmer, W., *How to publish DNA sequences with copyright protection.,* 20 NATURE BIOTECH. 217 (2002) (proposing distribution of verbal representations of DNA compounds as mp3 files); In a discussion of the commercial and legal implications of the Human Genome Project, the Committee on Mapping and Sequencing the Human Genome of the National Research Council was seduced by the fact that human genes can be represented as a series of As, Ts, Cs, and Gs to assume that although any new materials developed during the project would be protected by patent, the potential intellectual property mechanism for protecting human genome sequences would be copyright. COMMITTEE ON MAPPING AND SEQUENCING THE HUMAN GENOME OF THE NATIONAL RESEARCH COUNCIL, MAPPING AND SEQUENCING THE HUMAN GENOME at 91 (1988) (Should it be possible to copyright sequences from the human genome and if so, by whom? . . . This committee believes that human genome sequences should be a public trust and therefore should not be subject to copyright.")
[179] "Literary works" are defined by the Copyright Act as "works . . . expressed in words, numbers or other verbal or numerical symbols or indicia, regardless of the nature of the material objects, such as books, periodicals, manuscripts, phonorecords, film, tapes, disks, or cards, in which they are embodied." 17 U.S.C. § 101.

decipher code is a literary work.[180] The series of 0s and 1s that represent a computer program is a literary work. There seems no reason to exclude a series of As, Ts, Cs, and Gs from the definition of literary work. It easily complies with the single requirement set forth in the definition of "literary works" that it be expressed in verbal symbols or indicia.[181]

But if we conclude that copyright subsists in a literary work that is a series of As, Ts, Cs, and Gs, what does this mean for copyright and DNA? Is any series of the letters A, T, C and G a "DNA sequence?"[182] In his article discussing copyright protection for DNA, Andrew Torrance describes a yearly Mystery Hunt conducted by students from MIT. One year, students wrote a coded puzzle clue

---

[180] *See* Reiss v. Nat'l Quotation Bureau, Inc., 276 F. 717 (S.D.N.Y. 1921) (in which copyright was found in strings of letters which did not form words).

[181] It may be possible to argue that letters are components of words, and therefore not verbal symbols, but given the expansive language of the statute that indicates there are verbal symbols or indicia other than words, such an argument would be difficult to sustain.

[182] Discussions of DNA technology and copyright often pose the question "Are DNA sequences copyrightable?" *See* Dan L. Burk, DNA Copyright in the Administrative State, 51 UC Davis L. Rev. 1297, 1299 (2018) ("For nearly three decades, academics have toyed off and on with the question of copyright protection for recombinant DNA sequences."); Stephen R. Wilson, *Copyright Protection for DNA Sequences: Can the Biotech Industry Harmonize Science with Song*, 44 JURIMETRICS 409, 423 (2004) (discussing "attain[ing] copyright protection for DNA sequences by transforming them into digital music files"); Andrew W. Torrance, *DNA Copyright*, 46 VAL. U. L. REV. 1, 29 (2011) ("DNA sequences should be eligible for copyright protection."); Andrew Torrance, Synthesizing Law for Synthetic Biology, 11 MINN. J. L., SCI. AND TECH. 629, 648 (2010) (discussing consequences for the "copyrightability of synthetic DNA sequences"); Christopher M. Holman, Claes Gustafsson & Andrew W. Torrance, *Are Engineered Genetic Sequences Copyrightable?: The U.S. Copyright Office Addresses a Matter of First Impression,* 35 BIOTECH. L. REP. 103, 103 (2016) ("[T]he argument in favor of extending copyright to engineered DNA sequences has only gotten stronger. . .."). However, there are exceptions. *See,* Donna Smith, Comment, *Copyright Protection for the Intellectual Property Rights to Recombinant Deoxyribonucleic Acid: A Proposal*, 19 ST. MARY'S L.J. 1083, 1099 (1988) ("Copyright appears to be a viable alternative for the protection of intellectual property rights to rDNA molecules."); Andrew W. Torrance, *DNA Copyright,* 46 VAL. U. L. REV. 1, 35 (2011) ("DNA molecules are copyrightable. . .."). The term "DNA sequence" might have relevance in scientific communications as used to refer to the order of nucleotides in a DNA compound. *See, e.g.*, Assoc. for Molecular Pathology v. U.S. Patent and Trademark Office, 702 F. Supp. 2d 181, 194 (S.D.N.Y. 2010), aff'd in part & Rev'd in part, 689 F.3d 1303 (Fed. Cir. 2012, aff'd in part & Rev'd in part, 133 S. Ct. 2107 (2013). (defining "nucleotide sequence" as the "linear order of DNA nucleotides that make up a polynucleotide, such as a gene. . ..").

using a series of As, Ts, Cs, and Gs.[183] In order to interpret the clue, it was necessary to apply the rules of the genetic code to convert the series of As, Ts, Cs, and Gs into the letters representing the 20 amino acids in proteins. The resulting letters spelled out the clue in English. Some commentators have implied that the copyrightability of the clue has implications for the copyrightability of DNA.[184] Indeed, at the end of his recitation of the MIT game, Andrew Torrance concludes that even "DNA sequences" that serve the more traditional purpose of participating in the cellular process of making proteins may qualify for copyright protection "to the extent that function does not dictate structure, and expression is not unduly constrained."[185] But, a clue written in As, Ts, Cs, and Gs is not a "DNA sequence." It is simply a game clue written in code. This puzzle clue is copyrightable to the same extent any other coded message would be,[186] and the fact that it may be copyrightable means nothing for the copyrightability of DNA compounds generally. Granting copyright protection to a code written to be deciphered using the rules of the genetic code says nothing about the copyrightability of DNA. Imprecisely labelling the As, Ts, Cs, and Gs that make up the puzzle clue a "DNA sequence" implies that the series should be treated under the copyright law differently than other literary works, and that the copyrightability of this clue has some impact on the copyrightability of DNA compounds in general.[187]

It should be clear from this example that the letters A, T, C and G can represent any number of things. Conversely, any four letters or symbols could be used to represent a DNA compound in a writing or in a digital database. There is no rule of construction that excludes certain sequences of four letters from representing DNA compounds. Nor is there a rule of construction that defines certain sequences of four letters as necessarily representing DNA compounds. The

---

[183] SHOTGUN WEDDING (last visited Feb. 10, 2021), http://web.mit.edu/puzzle/www/2005/setec/shotgun_wedding/

[184] Andrew Torrance states that this "DNA sequence" would be "readily eligible for copyright protection" due to it "having little or no functionality and abundant expression" and implies that this has some relevance to the copyrightability of DNA generally. Andrew W. Torrance, *DNA Copyright*, 46 VAL. U. L. REV. 1, 3, 36 (2011); *See also* Pamela Samuelson, Evolving Conceptions of Copyright, 78 U. Pitt. L. REV. 17, 85 (2016).

[185] Andrew W. Torrance, *DNA Copyright*, 46 VAL. U. L. REV. 1, 3 (2011).

[186] On different grounds,

[187] Andrew W. Torrance, *DNA Copyright*, 46 VAL. U. L. REV. 1, 36 (2011).

only circumstance in which copyright protection to a series of As, Ts, Cs and Gs has any relevance to copyright protection for DNA technology is if that series of As, Ts, Cs and Gs represents a DNA compound.

So, let's consider copyright protection for a literary work consisting of an original series of As, Ts, Cs and Gs that is a verbal representation of a DNA compound. A written series of As, Ts, Cs, and Gs representing the sequence of nucleotides in a DNA compound certainly meets the minimal statutory requirements of a "literary work" in the same way that the 0s and 1s of computer object code and nonsense words[188] meet that requirement. However, while it may be a literary work that describes a new chemical compound, describing a new physical entity does not make it a new type of literary work. Copyright protection for such a work should be subject to the limitations that apply to any other copyrightable work. Most relevant to information technologies such as DNA, and perhaps most confounding, are the limitations set forth in Section 102(b) of the Copyright Act. Section 102(b) provides, "[i]n 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."[189]

For the purpose of copyright law, there are two ways to metaphysically conceive of literary works that are representations of DNA compounds. The first is to consider the literary work to be separate and apart from the underlying information stored in the DNA compound. In this formulation, the literary work *describes* the DNA compound. The second is to consider the literary work as an alternative manifestation of the DNA compound. In this formulation, the literary work functions in the same way as the DNA compound to store the information that is in the compound. At first blush, it seems ridiculous to equate a series of As, Ts, Cs and Gs with a DNA compound. However, evidence from copyright jurisprudence in the context of computer technology points to the second formulation as being more convincing to courts. With either formulation, as long as one stays on the path, either fork leads to the same destination, a place where

[188] *Reiss v. National Quotation Bureau, Inc.*, 276 F. 717 (S.D. NY 1921).
[189] 17 U.S.C. § 102(b).

copyright protection does not extend to functional genetic DNA.

B. The Scope of Copyright for Verbal Representations as Descriptions of DNA Compounds

Under the first formulation, a verbal representation of a DNA compound is a literary work which describes the DNA compound. Under well settled doctrine concerning copyright in literary works, it is the literary work itself, or in other words, "the language that an author uses to explain, describe, or express whatever ideas or useful arts she may have discovered or created that copyright protects. . . ."[190] Thus, although Einstein's articles laying out the special and general theories of relativity were copyrightable literary works, copyright protection did not extend to the core equations, such as the famous $E=mc^2$.[191] Similarly, even if a verbal representation describing a DNA compound were found to be original enough to garner copyright protection, no exclusive rights to the described DNA compound would be granted by copyright. To use a visual analogy, a lithograph of a paint can does not grant protection to the paint can itself. For example, the artist, Wayne Thiebaud, has created several paintings and lithographs depicting cans of paint. The copyright that subsists in these works of art does not prohibit someone from manufacturing the cans depicted.

An example may be useful. A recent case in which Judge Seibel of the United States District Court for the Southern District of New York bent over backward to explain to a pro se plaintiff what copyright does and does not protect provides an example of how this concept may be applied to the functional aspects of copyrightable works. In *Perry v. Mary Ann Liebert, Inc.*, the Court's decision explained why the defendant's figure depicting a modified metabolic pathway invented by the plaintiff did not infringe the plaintiff's own figure depicting the pathway. Judge Seibel relied on the differences between the colors and shapes used in the defendant's figure and those used in the plaintiff's figure to find noninfringement. The court found that the similarities between "Plaintiff's and

---

[190] *See*, Pamela Samuelson, *Why Copyright Law Excludes Systems and Processes*, 85 TEXAS L. REV. 1921, 1936 (2007). *See also*, H. R. Rep. No. 103-388, at 23 (1993) ("[A] certificate of registration on a scientific treatise would not extend to the formula therein, although it would extend to an original explanation of the formula.").

[191] Amer. Dental Ass'n v. Delta Dental, 126 F.3d 977, 979 (7th Cir. 1997).

Defendant's diagrams, few as they are, are the result of ***scientific*** fact[192] 'that is free for the taking,' not 'due to protected aesthetic expressions.'"[193] Without explicitly stating so, Judge Seibel defined the plaintiff's figure as a pictorial or graphic work describing the metabolic pathway, and applied the criteria, such as color or shape, often used to assess similarity between pictorial or graphic works. Imagine the reaction of the plaintiff, a biochemist PhD, when she learned that her diagram depicting the metabolic pathway that she altered to delay the effects of fruit ripening was not "substantially similar" to the defendant's diagram depicting the same altered pathway because one diagram uses thick black arrows while the other uses thin colorful arrows, or because one diagram uses boxes while the other uses boxes and ovals.[194]

In the case of verbal representations of DNA compounds, there may be minimal if any "protected aesthetic expression." One can imagine a similar reaction from a biochemist if told that a verbal representation of a genetic DNA compound written as AATTTGGCGGGTTT copied from another verbal representation of a genetic DNA compound written as AattTggCgggTtt would not be infringing. The first sequence may be copyrightable as a literary work just as the plaintiff's figure was copyrightable as a graphic work in *Perry*. However, the second sequence may not constitute copyright infringement because, although the idea or useful art created – the DNA compound – is lifted in whole from the original work, that element is unprotected by copyright. The capitalization may be protected expression just as the color of the arrows and the shapes of the boxes were protected expression in *Perry*. Because the second sequence did not copy the capitalization of the first, there is no infringement.

This conclusion may seem absurd to scientists who has put their hearts and souls into their scientific creations, but nothing in copyright law compels the conclusion that copyright will necessarily protect the commercially (or

---

[192] The plaintiff discovered that the modified metabolic pathway described in the article can be induced by introducing a certain chemical to plant cells. The diagrams displayed the introduction of the chemicals and the following reactions that take place in the cells. Therefore, I would argue that the diagrams depicted an invention rather than a fact. Perry v. Mary Ann Liebert, Inc., 2018 U.S. Dist. LEXIS 93513 at *2.

[193] Perry v. Mary Ann Liebert, Inc., 2018 U.S. Dist. LEXIS at *19.

[194] Perry v. Mary Ann Liebert, Inc., 2018 U.S. Dist. LEXIS at *19.

intellectually) valuable aspect of any work. Indeed, excluding the commercially valuable functional aspects of an information technology work from copyright protection should be much easier to swallow than excluding the commercially valuable facts in a compilation from copyright protection. Despite the exclusion of facts from the copyright protection afforded compilations, both courts and commentators, perhaps out of a sense of equity, are sometimes inclined to extend copyright protection to the commercially valuable (and costly to gather) information disclosed in those compilations.[195] However, information technology works, as opposed to informational works, present no such quandary. Functional aspects of an information technology work are often already protected by intellectual property law, specifically patents, as well as trade secrecy even if they are excluded from copyright protection.

In many ways, copyright protection for a series of As, Ts, Cs, and Gs defined as a literary work describing a DNA compound can be compared to a photograph of some paint on a board. Both the DNA compound and the paint on a board store a work which may be copyrightable or not. The DNA compound may store a motion picture or a process of synthesizing protein. Paint on a board may store a work of art or a road sign. The sequence of As, Ts, Cs, and Gs is an accurate description of the DNA compound with that sequence of nucleotides[196] just as a photograph may be an accurate description of the paint on the board. As discussed in detail by Justin Hughes, a photograph that is simply an accurate representation of an uncopyrightable work (in my example, the road sign) "has no copyright

[195] Jane C. Ginsburg, *Creation and Commercial Value: Copyright Protection of Works of Information*, 90 Colum. L. Rev. 1865, 1875-80 (1990). Acknowledging this inclination, Jane Ginsburg proposed the explicit recognition of a two-tier copyright regime, with different scopes of protection for high authorship works, such as novels and narrative histories, and low authorship works, such as telephone directories and compilations of stock quotations, so called informational works. Jane C. Ginsburg, *Creation and Commercial Value: Copyright Protection of Works of Information*, 90 Colum. L. Rev. 1865, 1869 (1990).

[196] Back in 2001, Dan Burk argued that the series of As, Ts, Cs, and Gs "which seems to display the information" in DNA "is not by itself of interest" because much of the essential information of value is omitted. *See*, Dan L. Burk, *The Problem of Process in Biotechnology*, 43 HOUS. L. REV. 561, 586 (2006). However, now, in the era of mail order DNA and DNA printers, it is difficult to see how any information unique to a particular DNA compound is left out of the series of As, Ts, Cs, Gs that may be used to represent it.

protection at all."[197] Graphic representations that are merely "slavish copying" of automobiles[198] and photographs of transmissions parts[199] and spindle bearings[200] have been found to lack copyright protection.[201] Similarly, a verbal representation of a DNA compound that is simply an accurate description of an uncopyrightable creation, such as a genetic DNA compound, should have no copyright protection at all, or at least, copying the aspects that comprise the DNA compound should not be an infringement.[202]

Some scholars have taken a similar approach to assessing the appropriate copyright protection for computer code. Under that reasoning, computer code, a literary work which describes the compiled machine-readable fixation of a program, is a very accurate description of that software. As John Kidwell described, "If one conceives of a computer as an extraordinarily complicated set of electrical

---

[197] Justin Hughes, *The Photographer's Copyright – Photograph as Art, Photograph as Database*, 25 Harv. J. L. & Tech. 327, 361-64 (2012); *See also,* Cindy Alberts Carson, Laser Bones: Copyright Issues Raised by the Use of Information Technology in Archaeology, 10 HARV. J. L. & TECH. 281 (1997) (Concluding that medical and scientific imaging should similarly not be entitled to copyright protection.).

[198] Meshwerks, Inc, v. Toyota Motor Sales U.S.A., Inc. 528 F.3d 1258 (10th Cir. 2008).

[199] ATC Distrib. Grp., Inc. v. Whatever It takes Transmissions $ Parts, Inc., 402 F.3d 700 (6th Cir. 2005).

[200] J. Thomas Distrib.. V. Greenline Distrib., 100 F.3d 956 (6th Cir. 1996).

[201] *But see*, Tomelleri v. Zazzle, Inc. 2015 U.S. Dist. LEXIS 165007 (finding question of fact whether a scientifically accurate depictions of fish lack originality required for copyright).

[202] At times, courts jump to the infringement analysis before first identifying the work which is allegedly copied and making an assessment regarding its copyrightability. As noted by Michael Risch, "In all three instances in which [the Supreme Court] has rendered an opinion on whether copyright protection extends to a portion of a work, it has reached its decision by comparing the accused work with the copyright claimant's work, and not by issuing a declaration of uncopyrightability." Brief *Amicus Curiae* of Michael Risch in Support of Petitioner, Google LLC v. Oracle America, Inc., No. 18-956 (S. Ct. argued Oct. 7, 2020) (discussion of Perris v. Hexamer, 99 U.S. 674 (1879), Baker v. Selden, 101 U.S. 99 (1880) and Feist Publications, Inc. V. Rural Telephone Service Co., 499 U.S. 340 (1991)). Jane Ginsburg also discussed how the Second Circuit in Knitwaves, Inc. v. Lollytogs Ltd., "confused the question of copyright *scope* with its subsistence." Jane C. Ginsburg, *Creation and Commercial Value: Copyright Protection of Works of Information*, 90 Colum. L. Rev. 1865, 1897 (1990). *See also*, Horizon Comics Prods., Inc. v. Marvel Entm't, LLC, 246 F. Supp. 3d 937, 941 (S.D.N.Y. 2017) ("the Court looks to whether the alleged similarities are due to protected aesthetic expressions original to the allegedly infringed work, or whether the similarity is to something in the original that is free for the taking.").

switches and relays, . . . the entry of the program into the computer is nothing more than the translation of the description of the switch settings into the setting of the switches themselves."[203] The code may be copyrightable, but the program is not protected. As Lloyd Weinreb recognized, "[t]he representation of a program [or sequence of nucleotides in a genetic DNA compound] in code or some other symbolic form . . . may be copyrightable, to the extent that its concrete expression is original. The program [or sequence of nucleotides in a genetic DNA compound] that is represented, however, contains no expression and is not copyrightable. . . ."[204]

C. The Scope of Copyright for Verbal Representations as an Alternative Form in Which DNA Compounds Exist

Under the second metaphysical conception, the series of As, Ts, Cs, and Gs is not a description of the DNA compound, but rather, it stores the same information as the DNA compound itself. As recognized in the patent literature, "[a]lthough [verbal representations of] DNA sequences represent chemical compounds, they are more fundamentally carriers of information."[205] Rather than describing a DNA compound, a sequence of As, Ts, Cs, and Gs may be, in a sense, an alternative form in which the DNA compound exists. The sequence is an embodiment in another medium of the DNA compound with that sequence of nucleotides.[206] Most discussions about the copyrightability of computer programs have treated computer

[203] John A. Kidwell, *Software and Semiconductors: Why Are We Confused*, 70 MINN. L. REV. 533, 542 (1985).

[204] Lloyd Weinreb, *Copyright for Functional Expression*, 111 HARV. L. REV. 1149, 1168 (1998). I would alter Professor Weinreb's statement slight to add, that the program contains no *valuable* expression. A machine code program written 000111010101 has the same amount of nonfunctional expression as a DNA sequence AATTTGCG. A machine code program written 000111 01010101 may not infringe just as a DNA sequence AA TTT GCG may not infringe.

[205] Arti Rai, *Addressing the Patent Gold Rush: The Role of Deference to PTO Patent Denials*, 2 WASH. U. J. L & POL'Y 199, 204 (2000).

[206] Back in 2001, Dan Burk argued that the series of As, Ts, Cs, and Gs "which seems to display the information" in DNA "is not by itself of interest" because much of the essential information of value is omitted. *See*, Dan L. Burk, *The Problem of Process in Biotechnology*, 43 HOUS. L. REV. 561, 586 (2006). However, now, in the era of mail order DNA and DNA printers, it is difficult to see how any information unique to a particular DNA compound is left out of the series of As, Ts, Cs, Gs that may be used to represent it.

code in this manner. As Samuelson, Davis, Reichman, and Kapor stated, "source code is the medium in which a program is created."[207]

This formulation is probably compelling to scholars and judges in the context of computer programs because computer code can now be converted into the program stored as the two electrical states on a tape or other digital storage device through procedures that function in a black box and without intervention on the part of a person.[208] As recognized by Kidwell, when a program is loaded into a computer, the description (or in other words, the program code) of the electrical switch settings "at a certain point become the switch settings [or in other words, the program contained in software]."[209] In effect, to someone who cannot see in the box, the program code appears to be an alternative medium in which to store the information stored in the program. Similarly, a verbal representation of a DNA compound can be converted into the DNA compound itself through procedures that function in a black box and without intervention on the part of a person. The verbal representation of a DNA compound appears to be an alternative medium in which to store the information stored in the DNA compound.

If the information stored in a DNA compound is a copyrightable work, such as a novel or motion picture, this alternative conception of the verbal representation leads to a satisfying result. Copyright may prohibit copying the series of As, Ts, Cs and Gs verbally representing the compound. Section 102(a) of the Copyright Act states that copyright subsists, in 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." A novel or motion picture fixed in a DNA compound is perceived with the aid two machines. First a sequencing machine identifies the sequence of

---

[207] *See* Pamela Samuelson, Randall Davis, Mitchell Kapor and J.H. Reichman, *A Manifesto Concerning the Legal Protection of Computer Programs*, 94 COLUM. L. REV. 2308, 2316 (1994). ("The view of programs as texts has been widely adopted in the legal community."). In other words, computer code is considered to *be* the computer program contained in software, which stores information consisting of steps of a process performed by a computer. To apply Professor Weinreb's comparison of computer programs to gears in a car,[207] the program code becomes the gears rather than the description of the gears.

[208] In the era of computer programming using punch cards, this was not the case.

[209] John A. Kidwell, *Software and Semiconductors: Why Are We Confused*, 70 MINN. L. REV. 533, 542 (1985).

nucleotides in the DNA compound. At least with current technology, the sequence of nucleotides is then converted to digital storage with the aid of a computer and then converted by the computer to the novel or motion picture perceivable by humans. Inputting a series of As, Ts, Cs and Gs representing the sequence of nucleotides into a computer simply circumvents the first step of sequencing the DNA compound. Therefore, a novel or motion picture fixed in a series of As, Ts, Cs and Gs is simply a shortcut that allows the novel or motion picture to be perceived with the aid of one fewer machine.[210]

But conceiving of verbal representations of DNA compounds as alternative embodiments of the compounds themselves leads to some absurd results when applied to genetic DNA. Under this formulation, if the DNA compound is a genetic DNA compound that stores the process for protein synthesis, the verbal representation of that genetic DNA compound also stores the process for protein synthesis.[211] Treating a verbal representation of a DNA compound as storing the information contained in the chemical compound allows literary works to do something they have not previously done. While traditional literary works may describe functional creations such as machines or processes, literary works in these information technologies *can be* functional creations.[212] Just as a genetic DNA compound can be functional, so can the verbal representation.[213] To extend the visual analogy of a lithograph picturing a paint can, if the lithograph is a visual embodiment of the paint can in the same way as the series of letters is a verbal embodiment of the DNA compound, it would seem as if the lithograph itself can now function to contain the paint. Despite this absurdity, if information technology has enabled *literary works* to be both functional and expressive in the same way

[210] Given the capacity of motion pictures to be stored as literary works, it is unclear into which category of copyrightable work such a work would fall. This question exists with respect to digitally stored motion pictures as well.
[211] *See* Section III.C., *infra*.
[212] *See*, Pamela Samuelson, Randall Davis, Mitchell Kapor and J.H. Reichman, *A Manifesto Concerning the Legal Protection of Computer Programs*, 94 COLUM. L. REV. 2308, 2323 (1994). ("Program text is, thus, like steel and plastic, a medium in which other works can be created. A device built in the medium of steel or plastic, if sufficiently novel, is patentable; an original sculpture built of steel or plastic is copyrightable.").
[213] *See, e.g.*, M. Scott McBride, *Bioinformatics and Intellectual Property Protection*, 17 BERKLEY TECH. L. J. 1331, 1337 (2002) (Discussing how scientists use databases of nucleotide sequences to compare and assign biological functions to particular or characteristic sequences.).

that pictorial, graphic and sculptural works can be, literary works should similarly be "protected in form but not their . . . utilitarian aspects."[214] In contrast to literary works which describe something utilitarian, a literary work storing only information that is a process may be useful in and of itself and should therefore be excluded from copyright protection as a utilitarian creation.

D. (Mis)applying the Merger Doctrine

The merger doctrine as applied to traditional literary works mandates that if the expression present in the copyrighted work is one of a limited number of ways to express an idea, the expression "merges" with the idea and should not be protected by copyright.[215] If a series of As, Ts, Cs and Gs is treated as a traditional literary text, in other words, one that describes the underlying chemical compound, the expression is the series of letters, and the idea is the DNA compound. Under the merger doctrine, if there are only limited ways a functional genetic DNA compound can be described, the expression in the literary work may merge with the uncopyrightable idea and be excluded from copyright protection. In other words, the result is the same under the doctrine of merger as it is when the verbal representation of the genetic DNA compound is considered merely an accurate depiction of an uncopyrightable work.

But as discussed in Section IV.C., *infra*, in the context of the written representations of computer software or DNA compounds, the literary work is often not treated as describing the underlying idea of the computer program or the DNA compound's nucleotide sequence. It is treated as containing the same information as the software or compound. It becomes an alternative form in which the software or compound exists. Thus, the function of the computer program or series of nucleotides in the DNA compound becomes the function of the literary work. Much to the detriment of clarity in copyright law, when literary works entered the realm

---

[214] Mazer v. Stein, 347 U.S. 201, 218 (1954) (citation omitted).

[215] Pamela Samuelson, *Reconceptualizing Copyright's Merger Doctrine*, 63 J. COPYRIGHT SOC'Y U.S.A. 417, 417, 419-20 (2016). *See also*, Southco, Inc. v. Kanebridge Corp., 390 F.3d 276, dissent at 291 (3rd Cir. 2004) ("The merger doctrine is a variation or application of the idea/expression dichotomy.").

of functional creations, they brought along with them the doctrine of merger.[216]

In the context of computer software, because courts have treated computer code as an alternative embodiment rather than a description of the underlying computer program, they have considered the function of the computer program, rather than the program itself, to be the "idea" with which an expression in computer code may merge. Under this formulation of the merger doctrine as applied to information technology, a functional literary work which consists of the only way of performing a function "merges" with the "idea" of the function and is excluded from copyright protection.[217] So far, so good, but the perversity of the merger doctrine to determine the proper extent of copyright protection for functional aspects of a work is disclosed when the inverse is asserted.

Pamela Samuelson traces the source of the merger doctrine to the early computer software copyright decision *Apple v. Franklin.*[218] In that decision, while recognizing that "[m]any of the courts which have sought to draw the line between an idea and expression have found difficulty in articulating where it falls," the court concluded that the pragmatic and proper line of inquiry should be "whether the idea is capable of various modes of expression."[219] The court concluded that if other programs can perform the same function of a particular program, then that program is an expression of the idea and hence copyrightable."[220] Thus, a

---

[216] Professor Samuelson traces the source of the application of the merger doctrine in computer software cases to the case *Apple v. Franklin*. In her view, the extension of the merger doctrine to software copyright cases stems from a misinterpretation of Baker v. Selden as restating the distinction between abstract ideas and expression rather than establishing the exclusion of procedures, processes, systems and methods of operations from copyright protection. Pamela Samuelson, *Why Copyright Law Excludes Systems and Processes*, 85 TEXAS L. REV. 1921, 1974 (2007). *See also*, H. R. Rep. No. 103-388, at 23 (1993).

[217] If computer code was more sensibly treated as describing the computer program, the underlying idea would be the computer program rather than the *function* of the computer program. Computer code without comments or other nonfunctional elements would in most cases merge with the idea of the computer program.

[218] Pamela Samuelson, *Reconceptualizing Copyright's Merger Doctrine*, 63 J. COPYRIGHT SOC'Y 417, 419-20 (2016).

[219] Apple Computer, Inc. v. Franklin Computer Corp., 714 F.2d 1240, 1253 (9th Cir. 1983).

[220] Apple v. Franklin, 714 F.2d 1240, 1253 (9th Cir. 1983). (The court based this test on a statement by the Second Circuit that a plurality of copyrights may exist for a plurality of ways of expressing an idea. *See*, Dymow v. Bolton, 11 F. 2d 690, 691 (2d Cir. 1926)).

justification for the copyrightability of functional aspects of literary works was born.[221]

The inverse of the merger doctrine has also been discussed as a justification for copyright protection for the functional aspects of "DNA sequences." In his discussion of the copyrightability of "DNA sequences," Torrance recognizes that the functionality of DNA compounds, particularly genetic DNA compounds, may limit copyrightability of such compounds.[222] However, he goes on to apply the inverse of the merger doctrine. Just as the court in *Apple* labelled the function of the computer program represented in computer code to be the computer code's idea, he treats the function of the underlying DNA compound's nucleotide sequence represented in the literary work as the "idea" with which an expression may merge. Applying the same reasoning as the court in *Apple*, he concludes that "if multiple DNA sequences could produce the same [protein] with a particular function, then any one individual [DNA] sequence would likely have much stronger copyright protection."[223] He continues, "[a]s DNA sequences increase in length and complexity, [] their eligibility for copyright protection would grow in proportion to

[221] M. Kramer Mfg. Co. v. Andrews, 783 F.2d 421 (4th Cir. 1986) (Stating that the accepted test for distinguishing the "idea" from the "expression" in the computer area was formulated in *Apple*.); Apple Computer, Inc. v. Formula Int'l, Inc., 725 F.2d 521, 526 (9th Cir. 1983) (Evidence that numerous methods exist for writing the programs involved proves that as in *Apple v. Franklin*, Apple seeks to copyright only its particular set of instructions, not the underlying computer process.); Autoskills, Inc. v. Nat'l Educ. Support Sys., 793 F. Supp. 1557, 1564-67 (1992) aff'd 994 F. 2d 1476, (10th Cir. 1993) (Stating that the court in *Apple v. Franklin* rejected an interpretation of *Baker v. Selden* which would exclude functional works from copyright protection. To distinguish protectable expression from unprotectable idea in the context of computer programs, "courts have looked for evidence of other programs in the marketplace which perform the same functions as the copyrighted work without employing the same methodology."); Oracle Am., Inc. v. Google, Inc., 750 F.3d 1339, 1367 (Fed. Cir. 2014) (citing *Apple* with approval and finding that "an original work – even one that serves a function – is entitled to copyright protection as long as the author had multiple ways to express the underlying idea."); *See also*, discussion of the dispute between Lotus Dev. Corp. and Paperback Software Int'l in Lloyd Weinreb, *Copyright for Functional Expression*, 111 Harv. L. Rev. 1149, 1154-63 (1998).

[222] Andrew W. Torrance, *DNA Copyright*, 46 VAL. U. L. REV. 1, 3 (2011) ("[E]ven 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.").

[223] Andrew W. Torrance, *DNA Copyright*, 46 VAL. U. L. REV. 1, 34 (2011).

their potential to be expressed in multiple ways."[224]

Applying the merger doctrine to other mixed functional and nonfunctional works demonstrates the folly of this novel application of the merger doctrine to allow copyright protection for functional aspects of literary works in information technologies such as DNA and computer software. Outside of the context of computer programs, the merger doctrine does not operate to permit copyright protection for the functional aspects of a work when there is more than one way to achieve that function.[225] Imagine if it did. The accounting form in *Baker v. Selden*[226] would be copyrightable because there was more than one form which could be used to perform the accounting system and therefore the "expression" in the form does not merge with the function.[227] The bicycle rack in *Brandir* would not be functional because there is more than one way to provide a parking space for a bike.[228] Indeed, a mousetrap would be copyrightable because there are multiple ways to build a better mousetrap that all perform the same function.[229]

---

[224] Andrew W. Torrance, *DNA Copyright*, 46 VAL. U. L. REV. 1, 36 (2011). As with all other functional works, the number of ways that a function may be "expressed" depends entirely on how the function is defined. There are many more DNA compounds that store the information necessary to synthesize a hormone, any hormone, than there are DNA compounds that store the information necessary to synthesize human growth hormone, specifically. But no matter whether the process is defined broadly or narrowly, it is a process. Thus, whether you define the process as synthesizing a hormone or synthesizing human growth hormone, the information stored in the DNA compound is a process.

[225] Arguments supporting the merger doctrine often state the inquiry as determining whether there is more than one way to "express" that function. Query what it means to express a function. I suspect that the term "express" is used to make the application of the merger doctrine in functional works sound more similar to the traditional application of the merger doctrine in the idea/expression context. An expression may describe a process or method of operation, but an expression probably is not a process or method of operation.

[226] Although *Baker v. Selden* has been cited as establishing the idea/expression dichotomy, more convincing analyses conclude that it "contributed the system and other useful art exclusions to § 102(b)." *See,* Pamela Samuelson, *Why Copyright Law Excludes Systems and Processes*, 85 TEXAS L. REV. 1921, 1928-36 (2007); Lloyd Weinreb, Copyright for Functional Expression, 111 Harv. L. Rev. 1149, 1176 (1998).

[227] Indeed, Baker's form was not identical to Selden's form.

[228] Brandir, Int'l, Inc. v. Cascade Pacific Lumber Co., 834 F.2d 1142 (2d Cir. 1987).

[229] At least 4,400 patents have issued for mousetraps. Nicholas Jackson, *Mousetraps: A Symbol of the American Entrepreneurial Spirit*, THE ATL. (March 28, 2011),

A showing that there is only one way to express something and still achieve the author's functional goal may constitute evidence that the expression is functional, but the inverse is not necessarily true. Evidence that there are many versions of an expression, whether it is a mousetrap, program code, or the verbal representation of a DNA compound, that may achieve the same functional goal does not preclude the expression from being functional.

While the idea/expression dichotomy, including the merger corollary, can operate effectively to establish the proper bounds of copyright protection for literary works which *describe* functional creations, only an outright exclusion for systems and other useful arts can establish the proper bounds for literary works which *can be* functional creations. As Samuelson points out, software case law has highlighted the deficiencies of applying the idea/expression dichotomy to exclude copyright protection for functional aspects of literary works.[230] If verbal representations of DNA compounds can now *be* functional, the inverse of the merger doctrine should not be used to establish copyright protection for their functional aspects. Rather, functional aspects of verbal representations of DNA compounds, just as functional aspects of other works, should be excluded from copyright protection.

V. Conclusion

Resolving ownership rights for the information stored in DNA matters. As our society and economy become less dependent on physical materials, information is becoming the currency of our interactions. Researchers no longer need to transfer chemical material, such as DNA compounds, among themselves. Information, in the form of the sequence of nucleotides in a DNA compound, is sent between labs,

---

https://www.theatlantic.com/technology/archive/2011/03/mousetraps-a-symbol-of-the-american-entrepreneurial-spirit/70573/. The mousetrap example does beg the general copyrightability question in the extreme case of a Rube Goldberg mousetrap. *See*, *How to Get Rid of a Mouse*, RUBE GOLDBERG (last visited Feb. 11, 2021), https://www.rubegoldberg.com/artwork/how-to-get-rid-of-a-mouse-2/. However, in that case, the proper inquiry would be whether an aspect of the mousetrap was not part of the function of a mousetrap and therefore possibly copyrightable. The question is not whether there are multiple ways to "express" a mousetrap.

[230] Pamela Samuelson, *Why Copyright Law Excludes Systems and Processes*, 85 TEXAS L. REV. 1921, 1974 (2007).

"there to be re-synthesized and expressed as needed."[231] One can imagine a similar future with respect to physical objects such as the Stanford Bunny.[232] As the means of physical production become more widely accessible, rather than receiving a product through the mail, 3-D printing instructions will be sent over the internet or stored in the material used to print the object itself. Ownership of the information necessary to produce an object is becoming equivalent to ownership of the object itself.

If the law regarding copyright protection for functional literary works follows the current path for copyright protection in computer code, copyright protection for verbal representations of DNA compounds may include protection for the functional aspects of DNA itself. Failure to distinguish between computer code, computer programs and their functions along with a failure to recognize the difference between literary works that describe a functional entity and literary works which are themselves functional has led to copyright protection for functional aspects of computer software. A similar lack of clarity about the distinctions between DNA compounds, verbal representations of DNA compounds, and the functions of DNA compounds threatens to lead to the same result. Now is the time to set copyright protection for DNA on a different path – before there are the statements of a CONTU-like commission with which to contend, before there exists inconvenient statutory language to address, and before there are conflicting

---

[231] Robert Carlson, *Open Source Biology and its Impact on Industry*, IEEE SPECTRUM 15, 17 (May 2001), http://www.eng.ucy.ac.cy/cpitris/courses/ECE001/notes/IEEEarticles/Open-source%20Biology%20And%20Its%20Impact%20On%20Industry%20-%20May%202001.pdf.

[232] Researchers have stored 3D printing instructions for a bunny figurine in DNA embedded in the figurine, itself. *See* Julian Koch, Silvan Gantenbein, Kunal Masania, Wendelin J. Stark, Yaniv Erlich, Robert N. Grass, *A DNA storage architecture to create materials embedded with memory*, 38 NATURE BIOTECH. 39, 40 (2020). https://www.nature.com/articles/s41587-019-0356-z.epdf?sharing_token=c2Pos7WwuJ9MtMGj-qdFQNRgN0jAjWel9jnR3ZoTv0M4Woj1cE3OBfuw5I5lxno_c7GoY2-6n89GH-ivEpAEquXjCDjHA8AZdMxio1_l4363ezz9mt81f8Ux0-ThMicOcJ3jN17Y29Zjoyaxwr_igBrd3adSox6_-oH3cNq6DJ1ULcp4ByGA1x5klZvn7uokkqRdDbeKAWtoswVNzrPPjDH9ZvIMjCntNYKs1wFv1aqEveze5ycUK_kclvwA58FZVaUZfx68IiSptq24UNUNJR2zrvMR0Vwz_5wNAwNHnqY%3D&tracking_referrer=spectrum.ieee.org.

court cases to reconcile.

Attempts to establish the appropriate scope of copyright protection for DNA by categorizing "DNA sequences" as works within or outside of the categories of works Congress intended to include in the Copyright Act miss the fundamental point. Recent advances enabling DNA compounds to store anything from sonnets to motion pictures make clear what has been true all along. DNA compounds are not works at all. DNA is a medium in which works are fixed. Just as with any other mechanical, electrical, magnetic, or chemical tangible medium of expression "now known or later developed," whether copyright prohibits the production of an unauthorized copy depends entirely on the nature of the information fixed in the copy. A novel or a work of art stored in a DNA compound should be entitled to copyright protection to the same extent as a novel or work of art stored in any other medium. DNA compounds that participate in the cellular processes to construct proteins should be excluded from copyright protection as part of a procedure, process, system, or method of operation.

Verbal representations of DNA compounds are copyrightable to the same extent as any other literary work and with the same exclusions applicable to any other literary work. As with any other literary work, copyright protection for the literary work should not extend to a functional DNA compound described in the literary work. Even if considered an alternate embodiment of the DNA compound itself, verbal representations of DNA compounds, as functional works, should only be protected by copyright to the extent they are not functional. Finally, the doctrine of merger should not be misapplied to functional works such as verbal representations of genetic DNA compounds to allow copyright protection for procedures, systems, processes, or methods of operation even if there are alternative ways to achieve the same function.

As patent rights in DNA compounds are limited by court decisions, there will inevitably be more discussion of copyright protection for DNA as inventors are motivated to turn to copyright to gain monopoly rights for DNA-based technology.[233] Extension of the copyright term to 70-plus years further incentivizes

[233] *See e.g.*, Devdatta Malshe, *Copyrighting DNA: An Off-Label Use,* 19 WAKE FOREST J. OF BUS. AND INTELL. PROP. L. 34, 37 (2018) (Predicting that the resulting action from the Supreme Court's *Myriad* decision "is now going to be a scramble to get man-made DNA copyright protection.");

those seeking exclusive rights to biological compounds to turn to copyright rather than patents to obtain those rights. Copyright protection for functional DNA could, at the discretion of the copyright holder, be used to create a commons of useful tools, but it seems foolish to rely on the goodwill of copyright owners to guarantee that what should not be protected by copyright remains free for the public to use. As Drew Endy, a scientist working in the new field of bioengineering has noted, copyright protection for functional DNA would be "horrifying" and "really dangerous if you mess it up" because copyright's lengthy term means that such exclusive rights "never end."[234]

Andrew W. Torrance and Linda J. Kahl, Bringing Standards to Life: Synthetic Biology Standards and Intellectual Property, SANTA CLARA HIGH TECH. L. REV. 227 (2014) ("now that natural-source DNA molecules have lost their eligibility for patent protection, copyright stands ready to provide an existing alternative form of protection.").

[234] Christopher M. Holman, *Developments in Synthetic Biology Are Altering the IP Imperatives of Biotechnology*, 17 VAND. J. ENT. & TECH. L. 385, 459 (2015) (quoting Drew Endy, Professor, Stanford Univ., Keynote Address at the Stanford Law School Conference on Intellectual Property Law and the Biosciences 50:30-52:00 (Apr. 27, 2012), http://www.youtube.com/watch?v=Qku3OQ5O_U4