

Regulating Information in Molecules: The Convention on Biological Diversity and Digital Sequence Information

Charles Lawson

Griffith University, Australia

Abstract

The United Nations *Convention on Biological Diversity* and its subsequent *Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity* provide a framework to conserve biological diversity, sustainably use biodiversity components and fairly and equitably share their benefits. There is unresolved contention about treating information as a derivative of biological materials and a distinct commodity with a value that can be translated into definable benefits. This article addresses whether there is information *in* DNA sequences, finding that there is causal information but no intentional or semantic information, although the causal contribution remains difficult to determine. This article concludes that caution should be exercised in limiting access to information *in* DNA through regulation because of the perverse outcomes controlling potential uses and reducing incentives for others to use information in new and innovative ways.

Keywords: *Convention on Biological Diversity*; genetic resources; access and benefit-sharing; DNA sequences; digital sequence information; information.

Introduction

The United Nations *Convention on Biological Diversity* (CBD) and its subsequent *Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity* (Nagoya Protocol) proposed a framework to conserve biological diversity, sustainably use the components of biodiversity and fairly and equitably share the benefits from utilising genetic resources.¹ The basic scheme for fairly and equitably sharing the benefits from utilising genetic resources obliges Contracting Parties to the CBD and Parties to the Nagoya Protocol to consider implementing legislative, administrative and policy measures facilitating access to ‘genetic resources’ within their sovereign control with prior informed consent and mutually agreed terms (known as access and benefit-sharing (ABS)).² In this context, ‘genetic resources’ are defined as ‘genetic material of actual or potential value’ and ‘genetic material’ as ‘any material of plant, animal, microbial or other origin containing functional units of heredity’.³ In practice, however, the term has a very flexible meaning, and Contracting Parties implementing the CBD may apply the term broadly to include most biological materials and derivatives.⁴ The Nagoya Protocol extends these obligations to include ‘derivatives’⁵ and to

¹ CBD, Art. 1; Nagoya Protocol, Art. 1.

² CBD, Art. 15; Nagoya Protocol, Arts. 5 and 6.

³ CBD, Art. 2. See also UNEP/CBD/WG-ABS/7/2, [18] and Annex ([3]).

⁴ See UNEP/CBD/WG-ABS/9/INF/1. See also UNEP/CBD/COP/3/20, [35]–[37].

⁵ A derivative is ‘a naturally occurring biochemical compound resulting from the genetic expression or metabolism of biological or genetic resources’: Nagoya Protocol, Art. 2. The interpretation is complicated because the term ‘derivative’ is included in the definition of ‘biotechnology’, and that term is then included in the definition of ‘utilization of genetic resources’ that is engaged in the fair and equitable benefit-sharing of the ABS obligations: Arts. 2, 5.1 and 5.2.



‘Traditional Knowledge associated with genetic resources’.⁶ The CBD has attracted 196 Contracting Parties and the Nagoya Protocol 133 Parties. Most ABS schemes rely on a contractual arrangement between a resource holder and the party seeking access to the resource that incorporates the CBD and Nagoya Protocol obligations, including the sharing of monetary and non-monetary benefits.⁷

Despite almost three decades of operation, concerns remain about the likely potential for the CBD and Nagoya Protocol to deliver significant benefits.⁸ As a result of these concerns, there has been a resurgent interest in extracting benefits from utilising information associated with genetic resources (e.g., downloading a DNA sequence from a publicly available database) as another source of potentially significant ABS benefits. However, failing to include or extend ABS to genetic information could undermine the existing ABS scheme because information can be utilised without the ABS arrangements and specifically benefit-sharing that apply to physical genetic resources.⁹ At the CBD and Nagoya Protocol forums, these concerns were captured by the term ‘digital sequence information’ (DSI). This term is a ‘place holder, without prejudice to future consideration of alternative terms’.¹⁰ The core of the contention is the ways and merits of treating DSI as a derivative of the materials within the ABS transaction itself, which becomes a distinct commodity with a value that the ABS scheme attempts to translate into definable benefits.¹¹

The purpose of this article is to address the concern that informational language in the context of genetic resources (e.g., ‘transcription’, ‘translation’, ‘coding’, ‘editing’, ‘proofreading’, ‘copying’, ‘gene expression’, ‘signals’, ‘program’ and ‘book of life’)¹² relating to morphological development and evolution potentially falling within the scope of the CBD and the Nagoya Protocol’s ABS arrangements is essentially flawed. While these are not new concerns in the discussion on information metaphors in the biological sciences, they are poignant because a bottom-up account of genetics based on the information flowing from DNA sequences infers a significance for a sequence that it might not have. In contrast, a top-down account traces an observable phenomenon to the products that result from the DNA sequence and *other* relevant causes. Put simply, as this article will demonstrate, the informational language used to describe molecular biology¹³ ‘leads to a misleading picture of possible explanations in molecular biology’¹⁴ and has been so pervasive in common understandings of genetics that it has probably limited the perspectives of policymakers addressing genetic resources. If this is correct, then this article argues that founding a legislative, administrative and policy scheme on this misleading picture of bottom-up information flowing from DNA sequences may perpetuate perverse outcomes and (further)¹⁵ undermine the purpose and integrity of ABS schemes.

In addressing these matters, the article is structured as follows. The next part outlines the dimensions of the DSI issue in the CBD and Nagoya Protocol forums. The following part traces the developments of the use of informational language in genetics, detailing the current theoretical framework for information in philosophy and law and distinguishing between the ideals of classical and molecular genes so that simplistic and ultimately misleading conceptions of DNA sequence do not undermine the role and place of information in DNA sequences. The next part discusses the implications of these theoretical threads for the regulation of DSI in the context of ABS. The final part concludes that there already exists adequate potential in the current ABS scheme of CBD Contracting States and Nagoya Protocol Parties to implement ABS legislative, administrative and policy measures to regulate DSI as a genetic resource. Alternatively, those implementing existing national ABS can include terms and conditions as part of prior informed consent and mutually agreed terms addressing DSI (as some Contracting Parties have done already). While initiating such measures is possible, there are potential problems. The result will be a matrix of different laws, policies and practices among the CBD Contracting States and Nagoya Protocol Parties that will likely perpetuate perverse

⁶ Nagoya Protocol, Arts. 7 and 12.

⁷ See Young, “Drafting Successful ABS Contracts”; Humphries, “Survey of Access and Benefit-Sharing.”

⁸ See, for example, Gaffney, “Open Access to Genetic Sequence Data”; Williams, “Conservation Policy: Helping or Hindering”; Rourke, “Policy Opportunities”; Laird, “Rethink the Expansion of ABS”; Lawson, “Information as the Latest Site”; Neumann, “Global Biodiversity Research”; Smith, “Biological Control”; Kupferschmidt, “Biologist Raise Alarm”; Bockmann, “Brazil’s Government Attacks”; Nijar, “Implementation of the Nagoya ABS”; Prathapan, “When the Cure Kills”; Lawson, “The Future of Information”; Humphries, “A Tiered Approach.”

⁹ Lawson, “Information as the Latest Site,” 19–26.

¹⁰ CBD/DSI/AHTEG/2018/1/4, [25] and Annex ([1]).

¹¹ Lawson, “The Future of Information,” 104.

¹² See, for example, Colucci-D’amato, “End of the Central Dogma” (neurobiology); Maynard-Smith, “Concept of Information” (developmental biology); Cooper, “Central Dogma of Cell Biology” (cell biology). See also Kay, “Who Wrote the Book of Life?”; Judson, “The Eighth Day of Creation”; Kalmus, “Cybernetical Aspect.”

¹³ It broadly captures the explanation of biology and emphasises the minuteness of biological entities, although this is not an entirely clear episteme: see Kay, “The Molecular Vision of Life,” 4–6.

¹⁴ Sarkar, “Biological Information: A Skeptical Look,” 187. See also Falk, “Genetic Analysis: A History,” 175.

¹⁵ See Laird, “Rethink the Expansion of ABS”; Lawson, “Information as the Latest Site.”

outcomes by controlling the potential uses of information and reducing the incentives for users of genetic resources to apply information in new and innovative ways. Consequently, this is likely to undermine the conservation and sustainable use of biodiversity because dealing with complicated laws, policies and practices will reduce the potential benefit-sharing opportunities for the uses of DSI. A preferable outcome would be an efficient and effective multilateral agreement balancing access and benefit-sharing that avoids the misleading picture of bottom-up information flowing from DNA sequences.

DSI as a CBD and Nagoya Protocol Issue

Formally recorded concerns about ABS and DSI in the CBD and Nagoya Protocol forums emerged in 2016¹⁶ with the decision to establish an *Ad Hoc* Technical Expert Group on Digital Sequence Information on Genetic Resources (AHTEG-DSI).¹⁷ The AHTEG-DSI compiled and synthesised views about DSI and commissioned a fact-finding and scoping study to consider the concept and scope of DSI and how DSI was currently used.¹⁸ Significantly, the indicative and contextual information ‘that may be relevant to the utilization of genetic resources’ considered by the AHTEG-DSI included:

- (a) the nucleic acid sequence reads and the associated data
- (b) information on the sequence assembly, its annotation and genetic mapping. This information may describe whole genomes, individual genes or fragments thereof, barcodes, organelle genomes or single nucleotide polymorphisms
- (c) information on gene expression
- (d) data on macromolecules and cellular metabolites
- (e) information on ecological relationships and abiotic factors of the environment
- (f) information on function, such as behavioural data
- (g) structure, including morphological data and phenotype
- (h) information related to taxonomy
- (i) modalities of use.¹⁹

The AHTEG-DSI concluded that more discussion about terminology was required to find a balance that could accommodate scientific, technological, market and other changes and provide legal certainty.²⁰ Recognising a lack of consensus and common ground about the scope of the CBD and Nagoya Protocol and the likely consequences for DSI on benefit-sharing through technology transfer, partnerships and collaboration, information exchange and capacity development,²¹ the AHTEG-DSI continued together with an open-ended working group to develop modalities for sharing benefits from DSI²² and reported their findings in 2018.²³ Reflecting the lack of agreement, the mandate of the AHTEG-DSI was extended, and they commissioned various additional studies on information traceability, sequence databases and domestic legal, administrative and policy ABS measures addressing DSI and benefit-sharing.²⁴ The outcomes of the AHTEG-DSI were to be considered by the Open-Ended Intersessional Working Group to Support the Preparation of the Post-2020 Global Biodiversity Framework and at the next Conference of the Parties (COP) to the CBD and the Meeting of the Parties (MOP) to the Nagoya Protocol in 2020.²⁵ However, the COP and MOP were postponed because of the global coronavirus pandemic.

The AHTEG-DSI has been a rich source of detail on DSI through its call for submissions and the commissioned peer-reviewed studies. The original commissioned scoping study identified the diversity of terms for DSI, including ‘*resources in silico, genetic sequence data, genetic sequence information, digital sequence data, genetic information, dematerialized genetic resources, in silico utilization, information on nucleic acid sequences, nucleic acid information, and natural information*’ (emphasis in original).²⁶ The study noted that more discussion of the terminology was required. It applied a basic conception of DSI as the order of nucleotides in a sequence that might be stored in a computer: ‘[DSI] is primarily the product of sequencing

¹⁶ CBD/COP/13/25, [321] and Decision XIII/16, [1]; CBD/NP/MOP/2/13, [153] and Decision 2/14, [1].

¹⁷ CBD/COP/13/25, [321] and Decision XIII/16, [4]. See also CBD/NP/MOP/2/13, [153] and Decision 2/14, [5].

¹⁸ CBD/DSI/AHTEG/2018/1/4, [17]. See also CBD/DSI/AHTEG/2018/1/2; CBD/DSI/AHTEG/2018/1/3.

¹⁹ CBD/DSI/AHTEG/2018/1/4, [23] and Annex ([2]).

²⁰ CBD/DSI/AHTEG/2018/1/4, [23] and Annex ([12]).

²¹ A draft decision with entirely bracketed text: CBD/SBSTTA/22/12, [35] and Recommendation 22/1; CBD/DSI/AHTEG/2018/1/4, [23] and Annex ([20]). See also Watanabe, “The Nagoya Protocol: The Conundrum”; Hammond, “Discussions on Sequence Information”; Scott, “Workshop Report: Genetic,” 37–38.

²² CBD/SBSTTA/22/12, [18]–[19].

²³ CBD/COP/14/14, [258] and Decision 14/20, [8–[12]; CBD/NP/MOP/3/10, [162] and Decision 3/12.

²⁴ CBD/COP/14/14, [258] and Decision 14/20, [11]. See also CBD/DSI/AHTEG/2020/1/7, [1].

²⁵ CBD/COP/14/14, [258] and Decision 14/20, [12] and [237] and Decision 14/34, [2]. See also CBD/WG2020/3/4.

²⁶ CBD/DSI/AHTEG/2018/1/3, 19–20.

technologies that have become faster, cheaper, and more accurate in recent years. The aim of DNA sequencing is to determine the order in which each of the four DNA nucleotides is arranged in the molecule'.²⁷

More recently, and extending previous work,²⁸ another commissioned study considered the concept and scope of DSI and how DSI was currently used concluding that the proximity of the information to the underlying physical genetic resource provided a logical basis to group information that could comprise DSI as follows: 'Group 1 – Narrow: DNA and RNA'; 'Group 2 – Intermediate: (DNA and RNA) + proteins'; 'Group 3 – Intermediate: (DNA, RNA and proteins) + metabolites'; and, 'Group 4 – Broad: (DNA, RNA, protein, metabolites) + traditional knowledge, ecological interactions, [and so on]'.²⁹ This study framed its discussion around the information flows represented by the 'central dogma' (DNA to RNA to protein to metabolites).³⁰

The AHTEG-DSI also commissioned a combined study on databases and traceability of DSI³¹ that essentially limited their considerations to databases holding nucleotide sequence data and traceability in the core database infrastructure, the International Nucleotide Sequence Data Collaboration (INSDC).³² This included nucleic acid sequence reads (the sequence of nucleotides, such as CGAAAGACCGGC) and the associated data and information on the sequence assembly, annotation and genetic mapping.³³ They found that some of these databases also included 'subsidiary information', which was broadly defined as information on gene expression, data on macromolecules and cellular metabolites, information on ecological relationships and other environmental data, functional data (e.g., behavioural data), structural data (e.g., morphological data and phenotype) and taxonomy data.³⁴ The study concluded that the INSDC's use of accession numbers (unique identifiers) facilitated database governance and traceability and that, at least in theory, this was a feasible mechanism for tracing nucleotide sequence data from a country or origin to a benefit-sharing user.³⁵

The other AHTEG-DSI commissioned study on domestic measures addressing the commercial and non-commercial use of DSI and benefit-sharing found four kinds of legislative, administrative and policy measures: regulating DSI as a distinct object of ABS and separate from the physical genetic resources; regulating DSI as a part of the utilisation of physical genetic resources; regulating DSI by requiring benefit-sharing (but not access) to cover the uses of DSI; and regulating DSI through other measures, such as compliance-related measures and monitoring mechanisms.³⁶ The study found that some jurisdictions explicitly included 'DSI' language like 'genetic information', 'genetic heritage', 'intangible components', 'gene sequences', 'sequence information', 'information' and 'information of genetic origin', while others interpreted their existing ABS legislative, administrative and policy measures as including 'DSI', such as 'genetic resources', 'genetic material', 'biological resources', 'associated knowledge', 'information of genetic origin', 'research results' and 'derivative'. However, the distinction between explicit and interpretive coverage was not necessarily clear.³⁷ Where DSI was regulated as a distinct object of ABS and separate from the physical genetic resources, the ABS schemes extended broadly to include information associated with the genetic resources. For example, Malaysia's ABS laws apply broadly to 'biological resources', which include 'genetic resources, organisms, microorganisms, derivatives and parts of the genetic resources, organisms, microorganisms or derivatives', 'the populations and any other biotic component of an ecosystem with actual or potential use or value for humanity' and 'information relating to' these 'biological resources'. In addition, the definition of 'derivative' includes 'information in relation to derivatives'.³⁸ Kenya's ABS laws apply to 'access', which means 'obtaining, possessing and using genetic resources conserved, whether derived products and, where applicable, intangible components, for purposes of research, bio-prospecting, conservation, industrial application or commercial use', where 'intangible components' are 'any information held by persons that is associated with or regarding genetic resources'.³⁹ Others regulate DSI by requiring benefit-sharing (but not access) obligations to cover the uses of DSI. For example, under India's ABS laws, benefit-sharing obligations apply to

²⁷ CBD/DSI/AHTEG/2018/1/3, 23.

²⁸ CBD/COP/14/14, [258] and Decision 14/20, [11(a)].

²⁹ CBD/DSI/AHTEG/2020/1/3, Annex (p. 32).

³⁰ CBD/DSI/AHTEG/2020/1/3, Annex (pp. 10–12).

³¹ CBD/COP/14/14, [258] and Decision 14/20, [11(c)–(d)].

³² CBD/DSI/AHTEG/2020/1/4, Annex (p. 14). The report importantly notes that outside the INSDC databases, the 'majority of such databases, after they are established during the project funding phase, are minimally, if at all, maintained, meaning webpages are infrequently updated, functions become defunct, or new data and bioinformatics tools are not added' (p. 18).

³³ CBD/DSI/AHTEG/2020/1/4, Annex (p. 14).

³⁴ CBD/DSI/AHTEG/2020/1/4, Annex (p. 14).

³⁵ CBD/DSI/AHTEG/2020/1/4, 64–66.

³⁶ CBD/DSI/AHTEG/2020/1/5, Annex (pp. 9–11).

³⁷ CBD/DSI/AHTEG/2020/1/5, Annex (pp. 11–12).

³⁸ *Access to Biological Resources and Benefit-Sharing Act 2017* (Malaysia), s. 4.

³⁹ *Environmental Management and Co-ordination (Conservation on Biological Diversity and Resources, Access to Genetic Resources and Benefit Sharing) Regulations 2006* (Kenya), reg. 2.

‘biological resource occurring in India or knowledge associated thereto’ for ‘research or for commercial utilisation or for bio-survey and bio-utilisation’,⁴⁰ where ‘research’ means the ‘study or systematic investigation of any biological resource or technological application, that uses biological systems, living organisms or derivatives thereof to make or modify products or processes for any use’; ‘commercial utilisation’ means ‘end uses of biological resources for commercial utilisation such as drugs, industrial enzymes, food flavours, fragrance, cosmetics, emulsifiers, oleoresins, colours, extracts and genes used for improving crops and livestock through genetic intervention, but does not include conventional breeding or traditional practices in use in any agriculture, horticulture, poultry, dairy farming, animal husbandry or bee keeping’; and ‘bio-survey and bio-utilisation’ means the ‘survey or collection of species, subspecies, genes, components and extracts of biological resource for any purpose and includes characterisation, inventorisation and bioassay’.⁴¹ The Indian law also provides that ‘[n]o person shall, without the previous approval of the National Biodiversity Authority, transfer the results of any research relating to any biological resources occurring in, or obtained from, India’.⁴² While there is no consensus apparent in the existing practices about the best ways to regulate DSI, there are, as the examples demonstrate, various forms of genetic information already subject to regulation in implementing CBD and Nagoya Protocol-consistent ABS schemes.

In addition to these commissioned studies, submissions of views and information to clarify the concept of DSI and benefit-sharing arrangements from using DSI have been made by CBD Contracting Parties and others, including other governments, Indigenous Peoples and local communities, relevant organisations and stakeholders.⁴³ A range of responses have been submitted essentially in three groupings: those arguing that DSI should not be a part of the CBD and Nagoya Protocol;⁴⁴ those favouring some accommodation of DSI;⁴⁵ and those favouring or already including DSI in their legal, policy and administrative ABS arrangements.⁴⁶ As a broad generalisation, technologically rich CBD Contracting Parties favour DSI not being a part of the CBD and Nagoya Protocol, and technologically poor CBD Contracting Parties favour DSI being accommodated or included in the CBD and Nagoya Protocol. As a useful summary of the way forward, the recent Open-Ended Working on the Post-2020 Global Biodiversity Framework considered a typology of possible regulatory options (although traditional knowledge associated with genetic resources was not addressed):⁴⁷ ‘Option 0: Status Quo’, addressing DSI under the existing arrangements through domestic ABS laws, policies and processes; ‘Option 1: DSI Fully Integrated into the [CBD] and the Nagoya Protocol’, addressing DSI as a genetic resource under the CBD and Nagoya Protocol and as an obligation under those agreements and implemented in domestic ABS laws, policies and processes; ‘Option 2: Standard [Mutually Agreed Terms]’, addressing DSI through an obligation to share benefits from the uses of DSI without restricting access to DSI itself through some kind of agreement with standard terms and conditions; ‘Option 3: No [Prior Informed Consent], No [Material Transfer Agreement]’, addressing DSI by requiring a payment or contribution for access or use of the DSI into a multilateral fund without the need for prior informed consent or mutually agreed terms and ABS contracts; ‘Option 4: Enhanced Technical and Scientific Cooperation’, democratising access and use of DSI so that each country has the capacity and opportunity to access and use DSI; and ‘Option 5: No Benefit Sharing from DSI’, no mechanisms are proposed and there is no benefit-sharing from the use of DSI.⁴⁸

Returning to the AHTEG-DSI and the indicative and contextual information ‘that may be relevant to the utilization of genetic resources’,⁴⁹ the commissioned study essentially considered four kinds of information according to ‘the flow of information from a genetic resource, particularly the degree of biological processing and proximity to the underlying genetic resource, to provide a logical basis to group information that may comprise DSI’.⁵⁰ Underpinning this ‘logical basis to group information’ was a particular conception of information in genetics founded in an ideal of the ‘central dogma’,⁵¹ which expresses genetic information as ‘nucleotide sequence information associated with transcription’, ‘protein sequence’ information associated with translation, ‘information associated with transcription and translation’ and ‘metabolites and biochemical pathways, thus comprising information associated with transcription, translation and biosynthesis’ and ‘extends to behavioural data, information on ecological relationships and traditional knowledge, thus comprising information associated with transcription,

⁴⁰ *Biological Diversity Act 2002* (India), s. 3.

⁴¹ *Biological Diversity Act 2002* (India), s. 2.

⁴² *Biological Diversity Act 2002* (India), s. 4.

⁴³ CBD/COP/14/14, [258] and Decision 14/20, [9].

⁴⁴ Examples include Australia, Canada, Japan and Korea.

⁴⁵ Examples include the European Union and its Member States and Switzerland.

⁴⁶ Examples include the African Union Commission on behalf of the African Group, Brazil, Ethiopia, India, Iran, Madagascar and South Africa.

⁴⁷ CBD/WG2020/3/4, Annex II (p. 13).

⁴⁸ CBD/WG2020/3/4, Annex II (pp. 15–17).

⁴⁹ CBD/DSI/AHTEG/2018/1/4, [23] and Annex ([2]).

⁵⁰ CBD/DSI/AHTEG/2020/1/3, Annex (p. 32). See also CBD/WG2020/3/4, Annex I (p. 8).

⁵¹ CBD/DSI/AHTEG/2020/1/3, 10–15.

translation and biosynthesis, as well as downstream subsidiary information concerning interactions with other genetic resources and the environment as well as its utilization, among other subsidiary information'.⁵² Importantly, the study also addressed the 'degree of biological processing and proximity to the underlying genetic resource' to distinguish between 'data' and 'information', the latter information being processed data.⁵³ The issue of the broader concept of biological information is addressed next. However, first, it is important to make a distinction here between information *about* DNA sequences and information *in* DNA sequences.

Information *about* DNA sequences is the vast quantity of information produced, collected, stored, accessed, managed and manipulated, including the order of nucleotides in a sequence, how the sequencing was conducted, annotations and functional analysis. This information is the subject matter of the information sciences bioinformatics applied to genetics:

What makes biology an information science in this sense is not anything about the nature of genes, but the fact that contemporary biology works with vast bodies of data that the unaided human mind is incapable of processing effectively.⁵⁴

The CBD already has an extensive mechanism to address this information *about* DNA sequences that is independent of the ABS obligations.⁵⁵ Essentially, the CBD has a general obligation to promote the exchange of information on the 'results of technical, scientific and socio-economic research', 'training and surveying programmes', 'specialized knowledge' and '[I]ndigenous and traditional knowledge as such and in combination with the technologies ['relevant to the conservation and sustainable use of biological diversity or make use of genetic resources']⁵⁶ and 'where feasible, include repatriation of information'.⁵⁷ There is a clearing house mechanism 'to promote and facilitate technical and scientific cooperation'⁵⁸ realised through decentralised databases and websites (information hubs) and national government websites.⁵⁹ The Nagoya Protocol Access and Benefit Sharing Clearing House is a part of the CBD's clearing house mechanism and applies only to ABS arrangements and 'access to information made available by each Party relevant to the implementation of this [Nagoya] Protocol'.⁶⁰ The CBD's Clearing House Mechanism (including the Nagoya Protocol Access and Benefit Sharing Clearing House), linked sites and sites linked to those sites set out information *about* DNA sequences and genetic resources more broadly.⁶¹

In contrast, information *in* DNA sequences is 'a theoretical entity which exists in the genome and explains biological phenomena'.⁶² Information *in* DNA sequences is the information *in* genetic resources as opposed to the information *about* genetic resources.

DNA Sequences as Information

It is uncontroversial that DNA is a linear sequence of molecules that can be presented as syntactic information in the language of the genetic code.⁶³ Letters of the alphabet making words that are joined into sentences, paragraphs and chapters represent syntactic information in language (here English). Similarly, photographs, music and computer programs are all, or can be rendered into, linear sequences of syntactic information in language (i.e., 0s and 1s of binary code). The proposition here is that because linear sequences in the form of words in sentences, paragraphs and chapters (and also photographs, music and computer programs) are information, then, similarly, DNA molecules in a linear sequence with a code represent information.⁶⁴ The question then is whether DNA molecules can actually be information.

⁵² CBD/DSI/AHTEG/2020/1/3, 32. See also CBD/WG2020/3/4, Annex I (p. 8).

⁵³ CBD/DSI/AHTEG/2020/1/3, 42–43.

⁵⁴ Griffiths, "Genetics and Philosophy," 145.

⁵⁵ For an overview, see Lawson, "Information as the Latest Site," 19–26.

⁵⁶ CBD, Arts. 16.1 and 17.1.

⁵⁷ CBD, Art. 17.2.

⁵⁸ CBD, Art. 18.3.

⁵⁹ See UNEP/CBD/WGRI/5/3/Add.2; UNEP/CBD/COP/12/11.

⁶⁰ Nagoya Protocol, Art. 14. UNEP/CBD/COP/10/27, [103] and Annex (Art. 14(1); Decision X/1, pp. 85–109). See also UNEP/CBD/COP/12/6, [51]–[58]; UNEP/CBD/ICNP/3/6.

⁶¹ See, for example, the Australian Government's Department of Environment and Energy website that sets out information about national strategies for biodiversity conservation, regulation and links to other websites that hold research and publications about research that includes information about genetic resources: <http://www.environment.gov.au/biodiversity>.

⁶² Griffiths, "Genetics and Philosophy," 144–145.

⁶³ See Crick, "Central Dogma of Molecular Biology"; Crick, "On Protein Synthesis."

⁶⁴ See, for example, Barbieri, "Definitions of Information and Meaning"; Barbieri, "Life and Semiosis."

The ideal of a DNA sequence as information traces back to at least 1953⁶⁵ and the ‘central dogma’ that information flows from DNA to RNA to proteins but not the other way out of proteins.⁶⁶ According to this account, the organism’s genome accumulates information (through the mechanisms of evolution) for transmission to the next generations.⁶⁷ The organism itself is merely the reservoir and transmitter of information.⁶⁸ Taken literally, a DNA sequence as information means that the arrangements of As, Ts, Gs and Cs represent the raw data, and they are themselves information. This is consistent with the ideal of life as information and programmable Boolean switches.⁶⁹ This might appear intuitively correct given the explosion of bioinformatics as a technological discipline exploiting information.⁷⁰ Unfortunately, this notion overlooks the complexity of genetics and gives undue weight to a particular conception of genotype as the causative (or purposeful) explanation (as addressed in detail below). The trajectory of this debate is important because developing regulatory schemes based on a particular perspective or preference regarding an unsettled theoretical foundation is likely to lead to bad laws and unforeseen consequences. At the heart of this problem is finding a common understanding for the term ‘information’ in genetics and how this might be addressed by law because, very crudely, it is *not* clear whether ‘information’ is addressing these molecules literally or metaphorically.⁷¹

A good entry point into this ongoing debate, and an obviously very brief account,⁷² is Charles Darwin’s 1859 theory of natural selection that set aside the idea that species were immutable, static and designed by a god. Darwin instead introduced the idea that species had adapted to their environments over many generations.⁷³ While Darwin posited that evolution and inheritance were linked, he also accepted that he was unable to explain the mechanism by which traits were inherited.⁷⁴ In 1865, Gregor Mendel provided such an account, positing from his experiments with peas that dominant and recessive elements were inherited. He traced those elements through hybrids as different constitutions and groupings of elements (*‘Faktoren’*).⁷⁵ Mendel’s insight was to mediate the relationship between genotypes (*‘genes’*) and *their* phenotypes (*‘unit characters’*)⁷⁶ by assigning the unobservable genotype to a phenotype (i.e., traits, such as seed texture, seed colour, pollen texture and pollen colour) and tracking the ‘dominant’ and ‘recessive’ phenotypes across hybridising crosses.⁷⁷ In this sense, Mendel’s elements were necessary for his explanation to work,⁷⁸ and as such, the Mendelian gene (albeit Mendel was no Mendelian)⁷⁹ refers to the unit of inheritance (*‘Zellelemente’*)⁸⁰ that predicted the apparent characters across generations. At its most simple, the Mendelian gene is an account of a mechanism for observed phenotypes from sexual crosses. This account posits a ‘gene’ (coined by Wilhelm Johannsen)⁸¹ to be an undefined unit of inheritance transmitted across generations and links the phenotype to a genotype.⁸² Here the genotype was the speculated and inferred cause of the observed phenotypes, with no explanation of the material and instrumental manifestation of the Mendelian gene itself.⁸³ Importantly, however, ‘what were studied were character differences, not characters, and what explained them were differences in genes, not the genes themselves’.⁸⁴ As a unit of inheritance, the Mendelian gene was a theoretical explanation of two kinds: first, the heritable factors an offspring receives, one from each parent in sexual crosses, albeit not an observable entity but an explanation of the observations of segregation

⁶⁵ Watson, “Molecular Structure of Nucleic Acids,” 737: ‘It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.’

⁶⁶ Crick, “Central Dogma of Molecular Biology”; Crick, “On Protein Synthesis.”

⁶⁷ Maynard-Smith, “Concept of Information,” 185–192.

⁶⁸ Morange, “History of Molecular Biology,” 2.

⁶⁹ For an elegant account of this perspective, see Rosenberg, “Darwinian Reductionism,” 56–93.

⁷⁰ See, for example, Ranganathan, *Encyclopedia of Bioinformatics*.

⁷¹ Griffiths, “Genetics and Philosophy,” 146–147.

⁷² The history of genetics is well traversed. See, for example, Falk, “Genetic Analysis: A History”; Carlson, “Mendel’s Legacy”; Keller, “The Century of the Gene”; Morange, “History of Molecular Biology”.

⁷³ Darwin, “On the Origin of Species.” For an overview, see Bowler, “Evolution: The History of an Idea.”

⁷⁴ Darwin, “On the Origin of Species,” 19–20.

⁷⁵ Mendel, “Versuche über Pflanzenhybriden,” 42.

⁷⁶ For language developed by Wilhelm Johannsen, see Johannsen, “The Genotype Conception of Heredity.” See also Roll-Hansen, “Sources of Wilhelm Johannsen’s Genotype Theory.”

⁷⁷ This is the passing over of a more complicated and intriguing moment in the history of genetics: see Falk, “Mendel’s Impact.”

⁷⁸ Griffiths and Stotz, “Genetics and Philosophy,” 15.

⁷⁹ Olby, “Mendel no Mendelian?”

⁸⁰ Mendel, “Versuche über Pflanzenhybriden,” 42.

⁸¹ See Johannsen, “Elemente der exakten Ererblichkeitslehre.”

⁸² See Wanscher, “Analysis of Wilhelm Johannsen’s.”

⁸³ For an account of this perspective, see Waters, “Genes Made Molecular,” 169–174 and the references therein.

⁸⁴ Waters, “Genes Made Molecular,” 172.

and independent assortment (the classical gene); and second, the material and instrumental entity of heredity (the molecular gene).⁸⁵

Both forms of the Mendelian gene persist,⁸⁶ although modern genetic practitioners often conflate the two forms:

When molecular biologists focus on nucleotide sequences, they think of genes in molecular concept. But at earlier stages of investigation, when they have not gotten close to specifying nucleotide sequences, they tend to think of genes in terms of the rougher-grained classical concept.⁸⁷

Rather than thinking of classical genes and molecular genes as separate theories,⁸⁸ most genetics practitioners consider a continuous theory addressed at two levels of resolution, with the classical genes being an ‘organic extension’ of the molecular gene.⁸⁹ Conceived this way, genetics is a reductionist, bottom-up account based in a physical sciences methodology using numerical analyses, with the outcome that the methodological and conceptual understanding of heredity is reduced to the sum of the physical and chemical properties of its building blocks, and that *is* DNA as the molecular *gene*, so the genotypes are the *causes* of phenotypes.⁹⁰ For the present purposes, it is significant that the later biochemical and molecular biological account of the material and instrumental manifestation of the gene overlooked this classical account of the units of a genetic and non-genetic context that resulted in the observed phenotype.⁹¹ Put simply, ‘molecular biologists *can now* determine the exact molecular identity of the relevant differences and explain how in general such differences produce phenotypic difference within a genetic context’ (emphasis added).⁹² This becomes clear when tracing the ideal of the material and instrumental manifestation of the molecular gene as opposed to the classical gene. This is important because the molecular gene has taken precedence as *the* account of genetics and gained a popular appeal⁹³ that overlooks much of the intriguing complexity and the role of other non-genetic (epigenetic)⁹⁴ factors in the observed phenotype, such as environmental effects.

The key moments in tracing the primacy of the material and instrumental manifestation of the molecular gene might be, hopefully uncontroversially,⁹⁵ stated as: Wilhelm Hofmeister first observed chromosomes in the 1840s;⁹⁶ Wilhelm Roux speculated that chromosomes are the carriers of inheritance in the 1880s;⁹⁷ Walter Sutton and Theodor Boveri correlated the action of chromosomes with the apparent results of Mendel;⁹⁸ Wilhelm Johannsen used the word ‘gene’ in 1909 to describe the fundamental physical and functional unit of heredity;⁹⁹ in 1910, Thomas Morgan proposed that genes (then called ‘factors’) are located on specific chromosomes;¹⁰⁰ Alfred Sturtevant provided a linkage map of genes in 1913;¹⁰¹ Theophilus Painter produced a cytological mapping of fruit fly salivary glands to localise genes at specific chromosome locations in 1934;¹⁰² George Beadle and Edward Tatum proposed the ‘one gene’ hypothesis based on their experiments showing that specific steps in metabolic

⁸⁵ For an engaging discussion of these different uses of the gene concept leading to Gregor Mendel being characterised as a methodological reductionist, the later re-discovery of Mendel’s work and Hugo de Vries applying the concept to the material and causal elements as a conceptual reductionist, see Falk, “Genetic Analysis: A History,” 4. See also Kitcher, “1953 and all That,” 336.

⁸⁶ See Carlson, “Mendel’s Legacy.”

⁸⁷ Waters, “Genes Made Molecular,” 183.

⁸⁸ See, for example, the different classical phenotypic marker gene-*P* and the molecular sequence gene-*D*: Moss, “What Genes Can’t.” See also Kitcher, “1953 and all That,” 336–337.

⁸⁹ Falk, “Genetic Analysis: A History,” 3–4.

⁹⁰ See Falk, “Mendel’s Impact,” 216–226. See also Fuerst, “The Role of Reductionism.”

⁹¹ See Falk, “Mendel’s Impact,” 229–233.

⁹² Waters, “Genes Made Molecular,” 183–184.

⁹³ See, for an example of a simplistic (and incorrect) representation of information flows: Wright, “DNA → RNA.”

⁹⁴ As used here and might be used in discussing molecular biology, ‘epigenetic’ means ‘the study of changes in gene expression that are mitotically heritable (via somatic cells) or meiotically heritable (via germ cells), and that do not entail changes in DNA sequence’: Griffiths, “Genetics and Philosophy,” 113.

⁹⁵ There are various accounts. See, for example, Portin, “The Evolving Definition of the Term Gene”; Allen, “Naturalists and Experimentalists”; Carlson, “The Gene: A Critical History.”

⁹⁶ Hofmeister, “Ueber die Entwicklung des Pollens”; Witty, “Pollen Development.” See also Kaplan, “The Genius of Wilhelm Hofmeister,” 1650–1652.

⁹⁷ Roux, “Ueber die Bedeutung der Kerntheilungsfiguren.” See also Hamburger, “Wilhelm Roux: Visionary,” 232–233.

⁹⁸ Sutton, “The Chromosomes in Heredity”; Boveri, “Über Mehrpolige Mitosen.” Noting that this claim is contested, see Martins, “Sutton and Boveri.”

⁹⁹ Johannsen, “Elemente der exakten Ererblichkeitslehre.” See also Johannsen, “The Genotype Conception of Heredity.”

¹⁰⁰ Morgan, “Sex-limited Inheritance,” 120–121. This proposition was more fully articulated in Morgan, “Mechanism of Mendelian Heredity.” Note, however, that Morgan appears to have been influenced by others: see Edwards, “Robert Heath Lock” (proposing that genes lying on a chromosome might account for linkages).

¹⁰¹ Sturtevant, “The Linear Arrangement.”

¹⁰² Painter, “A New Method,” 175.

pathways were disrupted by mutations in 1941;¹⁰³ Oswald Avery demonstrated that DNA was the hereditary material causing the heritable changes in 1944;¹⁰⁴ Erwin Chargaff clarified that the number of guanine and cytosine units and the number of adenine and thymine units were the same, hinting at the base pair makeup of the DNA;¹⁰⁵ James Watson and Francis Crick proposed the double helix structure for DNA in 1953;¹⁰⁶ Seymour Benzer proposed the conception of genes as linear structures along chromosomes (rather than being like beads on a necklace, they are instead divisible into smaller units of mutation and recombination) in 1955;¹⁰⁷ and, the elucidation of the genetic code by Crick and others in 1961.¹⁰⁸ The end of this track is the ideal of a linear molecular gene where the DNA sequence is considered the genotype and causative agent for the observed phenotype.

Intriguingly, Watson and Crick, in proposing the double helix structure for DNA in 1953, speculated about the ‘possible copying mechanism for the genetic material’¹⁰⁹ and ‘the precise sequence of the bases in the code that carries the genetic information’.¹¹⁰ This was essentially entrenching using information language in molecular biology that had started with terms such as ‘words’, ‘codes’, ‘messages’ and ‘texts’ in the 1930s and took solid hold in the 1940s.¹¹¹ Crick’s speculation later matured to the generalised rule for the informational transfer from one polymer to another (DNA to RNA, RNA to DNA, RNA to RNA, DNA to DNA and RNA to protein but *not* protein to protein,¹¹² protein to DNA and protein to RNA)¹¹³ so that ‘once “information” has passed into protein *it cannot get out again*’, where ‘information’ means ‘the *precise* determination of sequence, either of bases in the nucleic acid or of amino acid residues in the protein’ (emphasis in original).¹¹⁴ There are two parts to this claim that played out over the following decades. First, there is a coded sequence specificity between the DNA and the transcribed RNA and the translated polypeptide (sequence hypothesis).¹¹⁵ Second, the expression of the DNA sequence determines the RNA or a protein product such that *all* products are informed (specified or caused) by the DNA sequences (central dogma).¹¹⁶ Some of the details here matter: Crick accepted that protein synthesis involved ‘the flow of energy, the flow of matter, and the flow of information’, and his focus was the ‘information’.¹¹⁷ Crick later stated that:

it was abundantly clear by that time that a protein had a well-defined three-dimensional structure, and that its activity depended crucially on this structure, it was necessary to put the folding-up process on one side, and postulate that, by and large, the polypeptide chain folded itself up.¹¹⁸ This temporarily reduced the central problem from a three dimensional one to a one dimensional one ... The principal problem could then be stated as the formulation of the general rules for information transfer from one polymer with a defined alphabet to another.¹¹⁹

The key advances reinforcing these determinist and informational explanations of the now linear molecular gene, again hopefully uncontroversially, were: François Jacobs, Jacques Monod, Sydney Brenner, François Gros, Francis Crick and others’

¹⁰³ Beadle, “Genetic Control of Biochemical Reactions,” 505–506.

¹⁰⁴ Avery, “Studies on the Chemical Nature,” 149–150. See also Wyatt, “When Does Information Become Knowledge”; Hershey, “Independent Functions of Viral Protein.”

¹⁰⁵ Chargaff, “Chemical Specificity of Nucleic Acids.”

¹⁰⁶ Watson, “Molecular Structure of Nucleic Acids,” 737.

¹⁰⁷ Benzer, “Fine Structure of a Genetic Region,” 347.

¹⁰⁸ See Crick, “General Nature of the Genetic Code.”

¹⁰⁹ Watson, “Molecular Structure of Nucleic Acids,” 737.

¹¹⁰ Watson, “Genetical Implications,” 965.

¹¹¹ See Kay, “The Molecular Vision of Life,” 3–57.

¹¹² Crick did admit some ambiguity here, presciently considering prions: ‘There is, for example, the problem of the chemical nature of the agent of the disease scrapie’: Crick, “Central Dogma of Molecular Biology,” 562. Also note that there is evidence of proteins catalysing amino acid polymerisation: see, for example, Rout, “Prebiotic Template-directed Peptide Synthesis.”

¹¹³ Note that Crick asserts that reverse transcriptase using viral RNA as a template for DNA synthesis was not a reversal of the central dogma but, rather, a ‘misunderstanding’: Crick, “Central Dogma of Molecular Biology,” 561. See also Morange, “What History Tells Us.”

¹¹⁴ Crick, “On Protein Synthesis,” 153. See also Crick, “Central Dogma of Molecular Biology,” 562. See also Sarabhai, “Co-linearity of the Gene.”

¹¹⁵ ‘The Sequence Hypothesis ... in its simplest form assumes that the piece of nucleic acid is expressed solely by the sequence of its bases, sequence is a (simple) code for the amino acid sequence of a particular’: Crick, “On Protein Synthesis,” 152.

¹¹⁶ ‘The Central Dogma of molecular biology deals with the detailed residue-by-residue transfer of sequential information. It states that such information cannot be transferred from protein to either protein or nucleic acid’: Crick, “Central Dogma of Molecular Biology,” 561. For analyses of the various interpretations, see Camacho, “Central Dogma is Empirically Inadequate.” See also Thieffry, “Forty Years under the Central Dogma.”

¹¹⁷ Crick, “On Protein Synthesis,” 139–140.

¹¹⁸ This was later shown to be correct because the polypeptide chain folded itself into the thermodynamically most stable form. See Anfinsen, “The Kinetics of Formation.”

¹¹⁹ Crick, “Central Dogma of Molecular Biology,” 561.

discovery of mRNA in 1960;¹²⁰ the genetic code linking triplets of nucleotides (codons) to specific amino acids in 1961;¹²¹ Jacobs and Monod's explanation of a regulation mechanism (the *lac* operon) on a linear DNA molecule accounting for the relationship between DNA, RNA and proteins and pointing to a hierarchical network of regulation in 1962,¹²² Jim Shapiro and others' isolation of a bacterial gene (the *lac* operon) in 1969,¹²³ Howard Temin and David Baltimore's discovery of the enzyme that reversed transcription process making DNA from an RNA template in 1970,¹²⁴ David Jackson, Robert Symons and Paul Berg making recombinant DNA molecules in 1972,¹²⁵ Herbert Boyer and Stanley Cohen showing that engineered DNA molecules could be cloned in foreign cells in 1973,¹²⁶ and, then the sequencing of various genomes, including the bacteriophage ϕ X174 in 1978,¹²⁷ *Haemophilus influenzae* in 1995,¹²⁸ yeast *Saccharomyces cerevisiae* in 1996,¹²⁹ the nematode *Caenorhabditis elegans* in 1998,¹³⁰ the fruit fly *Drosophila melanogaster* in 1999¹³¹ and the human genome in 2000.¹³² The end of this track of enquiry was to cement the ideal of a linear molecular gene where the DNA sequence was popularly considered the genotype and causative agent for *all* the observed phenotypes.

The appeal of this approach was the focus on the simple explanatory power¹³³ of the information comprised by DNA both as a store of evolutionary accumulated changes and as a master plan for cell development and performance.¹³⁴ This was also the logic of a reductive physics and chemistry account where the molecule is the semantics of the transmitted information: 'the polypeptide phenotype is determined by the polynucleotide genotype'.¹³⁵ This has been, as the tracing of key advances above illustrates, amazingly heuristically successful. The key point, however, is that framed this way, the DNA sequence is conceived as a repository of meaning, aboutness and content (intentional or semantic information) regarding complex phenotypes so that the genotypes are privileged in the causes of phenotypes and the ideal that there is information *in* the DNA sequence.¹³⁶ The problem remains, however: If there is information *in* the DNA sequence, what kind of information is it?

Information in DNA Sequences

Even when Watson and Crick were proposing their DNA structure, the information metaphors were known to be problematic.¹³⁷ The postgenomic era following the release of the draft human genome sequence in 2001¹³⁸ revealed that the relationship between the 20,000–25,000 structural genes and the approximately 1,000,000 polypeptides of the proteome¹³⁹ was a lot more complicated than a mere linear sequence of nucleotides corresponding to a linear order of the gene products (RNA and amino acids in a polypeptide). Since then, the disruption of information flows has been repeatedly demonstrated, revealing that the DNA sequence *and* other factors (including non-genetic factors) are the contributors to RNA and protein sequence specificity (both quality and quantity): genes comprising complex regulatory networks affected by non-genetic factors result in a range of stable and robust phenotypes;¹⁴⁰ epigenetic acetylation, phosphorylation and methylation markings can regulate gene transcription;¹⁴¹ a range of *cis*- and *trans*-acting factors interact with the linear DNA sequence (e.g., transcription factors, promoters, activators, repressors, enhancers, silencers, and splicing factors), resulting in a diversity of RNAs (e.g., mRNA, rRNA, tRNA, lncRNA and RNAi) and protein forms/splice variants;¹⁴² reordering of the linear DNA sequences through

¹²⁰ See Morange, "History of Molecular Biology," 139–149 and the references therein.

¹²¹ Crick, "General Nature of the Genetic Code." See also Crick, "The Genetic Code – Yesterday."

¹²² Jacobs and Monod, "Genetic Regulatory Mechanisms."

¹²³ Shapiro, "Isolation of Pure *lac* Operon."

¹²⁴ Baltimore, "RNA-dependent DNA Polymerase"; Temin and Mizutani, "RNA-dependent DNA Polymerase."

¹²⁵ Jackson, "Biochemical Method for Inserting New Genetic Information."

¹²⁶ Cohen, "Construction of Biologically Functional Bacterial Plasmids."

¹²⁷ Sanger, "Nucleotide Sequence of Bacteriophage ϕ X174."

¹²⁸ Fleischmann, "Whole-Genome Sequencing."

¹²⁹ See Goffeau, "Life with 6000 Genes."

¹³⁰ The *C. elegans* Sequencing Consortium, "Genome Sequence of the Nematode *C. elegans*."

¹³¹ Adams, "The Genome Sequence of *Drosophila melanogaster*."

¹³² Venter, "The Sequence of the Human Genome"; Human Genome Sequencing Consortium, "Initial Sequencing and Analysis of the Human Genome."

¹³³ Compellingly 'simple and pedagogic' power that made molecular biology attractive: Morange, "History of Molecular Biology," 174.

¹³⁴ Maynard-Smith, "Concept of Information," 178–181. See also Schaffner, "The Watson-Crick Model."

¹³⁵ Falk, "Genetic Analysis: A History," 260.

¹³⁶ See Griffiths, "The Fearless Vampire Conservator," 175–198.

¹³⁷ See Keller, "Refiguring Life," 10; Kay, "Who Wrote the Book of Life?" 621–622. See also Fogle, "Information Metaphors."

¹³⁸ Venter, "The Sequence of the Human Genome."

¹³⁹ Muller, "Annotating the Human Proteome," 176.

¹⁴⁰ See, for example, Wagner, "The Role of Robustness in Phenotypic Adaptation."

¹⁴¹ See, for example, Cavalli, "Advances in Epigenetics."

¹⁴² See, for example, Morris, "The Rise of Regulatory RNA."

frameshifting, programme slippage or bypassing and codon redefinition,¹⁴³ and RNA editing resulting in a significantly larger transcriptome.¹⁴⁴ These examples confirm that there is not necessarily a consistent nexus between the DNA sequence as the sole source of information flowing from the DNA to an RNA and a protein.¹⁴⁵ Put simply, the evidence now clearly shows that there is not always a direct correspondence between the DNA sequence and protein because there is processing and modification of that sequence on the way from DNA to protein,¹⁴⁶ and modification of that processing and function can be from outside the DNA sequences (and particularly from the environment).¹⁴⁷ This calls for a nuanced depiction of DNA as information because a DNA sequence alone (genotype) *cannot* cause the whole organism. There are other contributing causes, such as the environment, and thus, the classical gene (observed phenotype) is not necessarily *only* caused by the molecular gene (deterministic and reduced to a DNA sequence). There are potentially better accounts of the information *in* DNA sequences, as argued in the following paragraphs.

A useful starting point is Shannon information theory, which posits a simple quantitative framework for describing correlations: two variables are correlated in some sense where the output of the channel depends on the input.¹⁴⁸ Here the correlation termed information has a special and technical mathematical sense that treats both sense and non-sense input messages as the same.¹⁴⁹ Taking this a step further, the information might be considered to have some content in the sense of natural signs and indicators¹⁵⁰ and that there is a correlation between a DNA sequence and an observed phenotype¹⁵¹ with information conveyed in the sense of its natural meaning.¹⁵² Thus:

The key currency in information theory is the entropy $H(X)$ of a random variable X . The entropy is a measure of uncertainty in the realization of X . If X takes on value X_i with probability p_i , the entropy $H(X) = \sum_i p_i \log p_i$. The key statistic in information theory is the mutual information $I(X;Y)$ between two random variables X and Y . The mutual information, defined as $I(X;Y) = H(X) - H(X|Y)$, measures how much we learn about the value of X by knowing Y .¹⁵³

If the DNA sequence is information—in this sense of information as the correspondence between the linear sequences of DNA nucleotides that specifies the linear sequences of RNA ribonucleotides and the linear sequences of amino acids (protein)—then Shannon information theory has some application with the input DNA sequence, the output amino acid sequence and the information transfer from DNA to amino acid modelled as the communication channel.¹⁵⁴ This accepts that DNA sequences do have a limited causal role¹⁵⁵ but not a broader intentional or semantic role.¹⁵⁶ Using Crick's words, this is 'the general rules for information transfer from one polymer with a defined alphabet to another',¹⁵⁷ where 'information was "merely a convenient shorthand for the underlying causal effect", namely the "precise determination of sequence"'.¹⁵⁸ This may also be conceived as a causal specificity between the DNA, the RNA and the coded amino acids,¹⁵⁹ representation of the RNA and amino acids in the DNA sequence and the correlation information between the DNA sequence, RNA and amino acids¹⁶⁰—the 'information that specifies the product is no longer carried in a three-dimensional structure but instead by the linear, one-dimensional order

¹⁴³ See, for example, Schmucker, "Dscam and DSCAM."

¹⁴⁴ See, for example, Tang, "Biological Significance of RNA Editing."

¹⁴⁵ Stotz, "With 'Genes' Like That," 905; Stotz, "Molecular Epigenesis," 542. For a different view see, for example, Baetu, "A Defense of Syntax," 713–714; Waters, "Genes Made Molecular," 174–182.

¹⁴⁶ See Falk, "Genetic Analysis: A History," 262.

¹⁴⁷ See Griffiths, "Genetic Information," 395–396.

¹⁴⁸ See Shannon, "Mathematical Theory of Communication." See also Fabris, "Shannon Information Theory."

¹⁴⁹ Shannon, "Mathematical Theory of Communication," 99.

¹⁵⁰ Dretske, "Epistemology and Information," 31.

¹⁵¹ See Godfrey-Smith, "Information in Biology," 104. Noting, however, that a sign, an object *and* an interpretant may actually be required for there to be any meaningful correlation. Hence, 'while tree rings might be a source of quantitative information, they would not mean anything about a tree's age unless there were an agent present who understood how tree rings are produced and how they relate to yearly growth': Kumar, "Information, Meaning, and Error," 91.

¹⁵² Grice, "Meaning."

¹⁵³ Bergstrom, "The Transmission Sense of Information," 161.

¹⁵⁴ See, for example, Bynum, "Informational Metaphysics," 205; Román-Roldán, "Application of Information Theory to DNA Sequence Analysis," 1188. Noting, of course, this might all be a misunderstanding: see Ben-Naim, "Entropy and Information Theory," 1185–1189.

¹⁵⁵ See, for example, Godfrey-Smith, "Information, Arbitrariness, and Selection," 204. For a contrary view, see Kitcher, "Battling the Undead" (arguing genetic coding has no explanatory weight).

¹⁵⁶ For other views, see Rosenberg, "Is Epigenetic Inheritance a Counterexample?"; Fantini, "Of Arrows and Flows"; Weber, "The Central Dogma as a Thesis."

¹⁵⁷ Crick, "Central Dogma of Molecular Biology," 561.

¹⁵⁸ Stotz, "Biological Information, Causality and Specificity," 369. See also Griffiths, "Genetic, Epigenetic and Exogenetic Information," 2.

¹⁵⁹ See Griffiths, "Measuring Causal Specificity"; Pocheville, "Comparing Causes," 93–102.

¹⁶⁰ See Shea, "Representation in the Genome," 314.

of elements in each sequence'.¹⁶¹ This might usefully be termed 'Crick information'.¹⁶² This Crick information also extends to regulatory DNA sequences (including non-coding sequences) and sequences of factors controlling transcription, splicing and editing and non-coding RNAs, where there is a correspondence between the DNA nucleotides and the RNA ribonucleotides and protein amino acids.¹⁶³ The next question, however, is whether DNA sequences can hold more information than just Crick information.

The main contribution from Mendel's genetics was to establish a process among scientists of an experimental tradition hybridising organisms and looking to draw inferences from the patterns of inheritance.¹⁶⁴ The quest for the causative agents shifted from the unobservable Mendelian gene to the molecular gene. This has privileged and preferred DNA sequence information with the presumption that DNA sequences are the cause of the subsequent transcription, translation and then protein action *and* the postgenomic events resulting in a phenotype that can be traced back to a causal DNA sequence. This is consistent with the early understanding of the central dogma that the linear order of nucleotides in DNA specifies the linear order of nucleotides in RNA and the linear order of amino acids in polypeptides (hence Crick information) and early analyses that identified DNA sequences as making the *actual* causal difference for the observed RNAs and proteins.¹⁶⁵ In addition to this is the regulation of linear genes. With the *lac* operon as the model in mind, for regulatory DNA sequences upstream of a DNA coding sequence for the observed RNAs and proteins, there was information about the conditions for expression and a programmed blueprint for the resulting cells, organs and organisms. The controversy is whether there is information *in* the DNA sequence that has a role in development (i.e., programs for individual organisms where the program carries the information for development) and evolution (i.e., inherited characters holding the accumulated information over evolutionary time). Putting this another way, DNA appears as a structure and mechanism with a purpose, function, end or goal as information for the programmed machine of the cell, organ, organism and so on, with mutation modifying the DNA and natural selection choosing fitter outcomes such that the DNA sequences direct the effects.¹⁶⁶ Information, in this sense, can generally be considered either causative or intentional or semantic.¹⁶⁷ This causative information, in addition to Crick information, is considered in this section, and the intentional or semantic information is considered in the next.

Causative Information

As set out above, causative information is about the quantity of information (i.e., order as opposed to disorder/entropy and not meaning) within a physical system so that information flows between a sender and a receiver through a channel about the correlations between the sender and receiver (hence Shannon information). The signal carries information about a source such that you can predict the state of the source from the signal:¹⁶⁸ 'whenever *Y* is correlated with *X*, we can say that *Y* carries information about *X*',¹⁶⁹ and, hence, 'the disease phenotypes carry information about disease genes',¹⁷⁰ 'genes carry information about phenotypes just as smoke carries information about fire'¹⁷¹ and 'genes contain information about the proteins they make, and ... about the whole-organism phenotype [in the same way as] there is an informational connection between smoke and fire, or between tree rings and a tree's age'.¹⁷² In the context of DNA sequences, the sequences can be the source, the whole organism the receiver and the channel conditions are the resources needed for the organism's life cycle.¹⁷³ For a causal information

¹⁶¹ Stotz, "Biological Information, Causality and Specificity," 368. Note that three-dimensional structures can now be reliably predicted: see Jumper, "Highly Accurate Protein Structure Prediction."

¹⁶² Griffiths, "Genetics and Philosophy," 40–41. See also Sarkar, "Genes Encode Information," 261–262, 281–284; Godfrey-Smith, "Information, Arbitrariness, and Selection," 203; Godfrey-Smith, "On the Theoretical Role of 'Genetic Coding'," 35; Godfrey-Smith, "Genes and Codes," 325. In more recent work, Paul Griffiths and Karola Stotz have 'reserved the term "Crick information" for a measure of the intrinsic information content of a sequence, rather than for the measure of the relationship between a sequence and its causes': Griffiths, "Genetic, Epigenetic and Exogenous Information," 7. See also Stotz, "Biological Information, Causality and Specificity," 367–370.

¹⁶³ Stotz, "Biological Information, Causality and Specificity," 380; Stotz, "With 'Genes' Like That," 538. Noting that this is not uncontroversial, see Sarkar, "Decoding 'Coding'-Information."

¹⁶⁴ Griffiths, "Genetics and Philosophy," 17. See also Falk, "Genetic Analysis: A History," 4.

¹⁶⁵ See, for example, Waters, "Causes that Make a Difference," 572–576. See also Rosenberg, "Is Epigenetic Inheritance a Counterexample."

¹⁶⁶ Maynard-Smith, "Concept of Information," 189–192. See also Griffiths, "Genetic Information," 398–401.

¹⁶⁷ Sterelny, "Sex and Death," 101. Note that "[i]f there is a relationship between intentional information and causal information it is a complex and distant one": Griffiths, "Genetic Information," 397.

¹⁶⁸ "Information" in the sense of "natural meaning": Grice, "Meaning."

¹⁶⁹ Bergstrom, "The Transmission Sense of Information," 189.

¹⁷⁰ Griffiths, "Genetic Information," 397. See also Griffiths, "Genetics and Philosophy," 159.

¹⁷¹ Bergstrom, "The Transmission Sense of Information," 161–162 and the references therein. See also Godfrey-Smith, "Information in Biology," 105–106.

¹⁷² Godfrey-Smith, "Information in Biology," 106.

¹⁷³ Griffiths, "Genetics and Philosophy," 159. See also Griffiths, "Proximate and Ultimate Information," 80.

account to be sufficient, holding the resources needed for the organism's life cycle constant would give information about the sequences.¹⁷⁴

As set out above, this causal information account is sufficient for Crick information because the DNA sequence is directly physically causally related to the RNA and amino acid sequence.¹⁷⁵ The DNA sequence CGAAAGACCGGC correlates (through the RNA sequence) with the amino acid sequence RKTG. Therefore, the DNA sequence CGAAAGACCGGC carries information about the amino acid sequence RKTG (Crick information). However, the causal information account is also sufficient where there is a correlation between a genotype and a phenotype *as well as* every other correlated non-genetic factor and the phenotype.¹⁷⁶ For example, 'an individual with the gene for achondroplasia will have short arms and legs [and] we can equally well say that a baby's environment carries information about growth; if it is malnourished, it will be underweight'.¹⁷⁷ In this sense, DNA sequence information (the gene for achondroplasia) is qualitatively the same as any other kind of correlated information (malnourished and underweight). Importantly, this information is not being used for an explanatory account to say what the DNA sequence product does or how it does it.¹⁷⁸ The apparent weakness of this causal account, so information in a richer sense, is that it makes information ubiquitous because it is not possible to distinguish genetic from non-genetic causes.¹⁷⁹ Put differently, DNA sequences do carry syntactic information (Shannon information) about RNA and amino acid sequences and some phenotypes,¹⁸⁰ but the DNA sequences cannot explain all the causal aspects of phenotypes, and every causal input, including non-genetic causes, will also be a source of information.¹⁸¹ This is the 'parity thesis' (or 'parity of reasoning') that the informational causal factors¹⁸² resulting in a phenotype are both the DNA sequences and other non-DNA elements, and as causal factors, the DNA sequences do not have any necessary priority or privilege.¹⁸³ The result is that without being able to weigh the different genetic and non-genetic causes, the best this causal information can be is about correlations, and they will be *any and all* correlations.

In attempts to bring more nuances to causal accounts, there have been further proposals. The starting point is accepting that phenotypes are the result of an interaction between DNA sequences (genes) and the environment (called the 'interactionist consensus')¹⁸⁴ and rejecting the ideal that DNA sequences (genes) have some necessarily superior causative role over the environment in determining a phenotype (called 'causal democracy').¹⁸⁵ In this sense, DNA sequences have a difference-making role when they actually affect the phenotype ('actual difference-makers'),¹⁸⁶ and this opens up the possibility of discriminating the contribution of a DNA sequence as 'causal relationships are relationships that are potentially exploitable for purposes of manipulation and control' from other genetic and environmental causes.¹⁸⁷ Using this ideal of 'causal specificity', it is then possible to distinguish the causal information from DNA sequences that provide information about their effects and quantify this information using Shannon information theory measures—effectively the reduction of uncertainty.¹⁸⁸ Essentially:

There is a causal relationship between variables *X* and *Y* if it is possible to manipulate the value of *Y* by intervening to change the value of *X*. 'Intervention' here is a technical notion with various restrictions. For example, changing a third variable *Z* that simultaneously changes *X* and *Y* does not count as 'intervening' on *X*. Causal relationships between variables

¹⁷⁴ See Griffiths, "Genetic Information," 398.

¹⁷⁵ See Šustar, "Crick's Notion of Genetic Information"; Griffiths, "Genetic, Epigenetic and Exogenetic Information," 1; Godfrey-Smith, "On the Theoretical Role of 'Genetic Coding'," 29. For other causal effects, see Weber, "Philosophy of Experimental Biology," 258–260.

¹⁷⁶ Griffiths, "Genetic Information," 398.

¹⁷⁷ Maynard-Smith, "Concept of Information," 189. See also Griffiths, "Genetic Information," 398.

¹⁷⁸ Levy, "Information in Biology: A Fictionalist Account," 642.

¹⁷⁹ Griffiths, "The Fearless Vampire Conservator," 183; Shea, "Inherited Representations," 2.

¹⁸⁰ This has become a dynamic field of study addressing measures of causal specificity. See, for example, Pocheville, "Biological Information as Choice"; Calcott, "Signals That Make a Difference"; Weber, "Discussion Note: Which Kind of Causal Specificity"; Pocheville, "Comparing Causes," 93–102; Griffiths, "Measuring Causal Specificity."

¹⁸¹ See Griffiths, "Three Ways to Misunderstand," 420–423. See also Stotz and Griffiths, "Biological Information, Causality and Specificity," 371–372.

¹⁸² Noting the different uses of "parity", see Stegmann, "Varieties of Parity," 913–915.

¹⁸³ Griffiths, "Three Ways to Misunderstand," 420–421 and the references therein. For counterviews, see, for example, Kjosavik, "Genes, Structuring Powers and the Flow"; Bergstrom, "The Transmission Sense of Information," 162; Neumann-Held, "Genes – Causes – Codes," 238–271. See also Shea, "Developmental Systems Theory."

¹⁸⁴ See Robert, "Embryology, Epigenesis and Evolution," 2. See also Kitcher, "In Mendel's Mirror," 290–300.

¹⁸⁵ See Kitcher, "In Mendel's Mirror," 290. See also Oyama, "Causal Democracy."

¹⁸⁶ Sterelny, "The Return of the Gene," 348–350. See also Waters, "Causes that Make a Difference," 572–576; Weber, "The Central Dogma as a Thesis," 599–605.

¹⁸⁷ Woodward, "Making Things Happen," 314.

¹⁸⁸ See Griffiths, "Measuring Causal Specificity." See also Griffiths, "Genetics and Philosophy," 78–84.

differ in how ‘invariant’ they are. Invariance is a measure of the range of values of X and Y across which the relationship between X and Y holds. But even relationships with very small ranges of invariance are causal relationships.¹⁸⁹

The idea here is that the more specific a causal relationship is between a DNA sequence and the observed phenotype, the more informational the DNA sequence will be: ‘The specificity of a causal variable is obtained by measuring how much mutual information interventions on that variable carry about the effect variable’.¹⁹⁰ As practical examples, RNA transcribed from DNA sequences in a cell involve the DNA sequence, RNA polymerase and several other proteins so that each is a cause of the resulting RNA molecule. However, it is only the DNA sequence that is a specific actual difference-maker because varying the DNA sequence varies the resulting RNA molecule.¹⁹¹ Extending this further, modelling measuring the mutual information between causes and effects of significant causes for DNA sequences and the simple production of *cis*-spliced mRNAs in a cell at a specific time showed the separate contributions of the DNA sequence and the *cis*-spliced mRNA variants.¹⁹² An example is the single *Drosophila melanogaster* Down syndrome cell adhesion molecule (DSCAM) coding sequence, with its 38,016 splice variant proteins and the same DSCAM coding sequence in *Homo sapiens* involving three splice variants.¹⁹³ The causes in this model are the DNA sequence and the *trans*-factors affecting the splicing so that the information between the RNA and splicing is the sum of the mutual information (causal specificity) between the DNA and RNA and between the RNA and splicing.¹⁹⁴ The contributions from the DNA sequence and the splicing can be decomposed so that a value can be assigned to the variation in RNA coming from the splicing and the number of splicing variants per DNA sequence.¹⁹⁵ The result is a value (in bits) that can be assigned to the contribution of the DNA sequence and the splicing so that their relative contributions (significant causes) can be assessed.¹⁹⁶ Thus, the single DSCAM coding sequence in *Drosophila* is 15.2 bits for the information coming from the splicing processes and 0 bits for the amount of information originating in the DNA sequence and preserved in splicing processes. Therefore, causation is entirely accounted for by post-transcriptional processing.¹⁹⁷ The same DSCAM coding sequence in humans is 1.6 bits for the information coming from the splicing processes and 1 bit for the amount of information originating in the DNA sequence and preserved in splicing processes. Therefore, causation is partly accounted for by the DNA sequence and partly by post-transcriptional processing.¹⁹⁸

This analysis shows that DNA sequences are not necessarily the most significant causes and that the significant causes need to accommodate the different spatial (i.e., cells, tissues and organisms) and temporal diversity because the measured contributions do change across spaces and times.¹⁹⁹ Thus, highly causally specific relationships are informational, and these occur for DNA sequences such that ‘[o]rganisms reproduce with a high degree of fidelity though the informational specificity of nucleic acids for proteins and functional RNAs’.²⁰⁰ These highly causally specific relationships, however, are also informational for *all* other biological systems and will apply to every fine-grained control over effects, such as antibody-mediated immune responses, enzymes for substrates and receptors for their ligands.²⁰¹ The consequence is that the contribution of the DNA sequence as a cause needs to be assessed for every instance, and only in some cases will the DNA sequence be the most significant cause and very, very rarely the *only* cause.

Intentional Information

In the early responses to the idea that it was not possible to distinguish genetic from non-genetic causes, there have been attempts to posit DNA sequence information as intentional or semantic.²⁰² Intentional or semantic information is the information of human thoughts and linguistic representations (utterances). It is intentional in that it is about the genotype of the phenotype that the genotype is intended to produce, although not necessarily the one *actually* produced.²⁰³ The key distinction between

¹⁸⁹ Griffiths, “Proximate and Ultimate Information,” 80. See also Griffiths, “Genetics and Philosophy,” 79–80.

¹⁹⁰ Griffiths, “Measuring Causal Specificity,” 538. See also Stotz, “Biological Information, Causality and Specificity,” 373.

¹⁹¹ Waters, “Causes that Make a Difference,” 573–574.

¹⁹² Griffiths, “Measuring Causal Specificity,” 543–550.

¹⁹³ Griffiths, “Measuring Causal Specificity,” 548–549.

¹⁹⁴ Griffiths, “Measuring Causal Specificity,” 548.

¹⁹⁵ Griffiths, “Measuring Causal Specificity,” 548.

¹⁹⁶ Griffiths, “Measuring Causal Specificity,” 549.

¹⁹⁷ Griffiths, “Measuring Causal Specificity,” 548–549.

¹⁹⁸ Griffiths, “Measuring Causal Specificity,” 548–549.

¹⁹⁹ Griffiths, “Measuring Causal Specificity,” 551.

²⁰⁰ Stotz, “Biological Information, Causality and Specificity,” 374.

²⁰¹ Stotz, “Biological Information, Causality and Specificity,” 374.

²⁰² For the key proponents, see Shea, “Inherited Representations”; Shea, “Representation in the Genome”; Maynard-Smith, “Concept of Information”; Sterelny, “The ‘Genetic Program’ Program”; Sterelny, “The Extended Replicator.”

²⁰³ Griffiths, “The Fearless Vampire Conservator,” 186. See also Godfrey-Smith, “Misinformation.”

intentional information and causative information is that intentional information need not be true²⁰⁴ (like ‘phlogiston or Pope Joan’).²⁰⁵ Molecular biology includes many terms that assert this intentionality, such as ‘messenger molecules’, ‘recognition sites’, ‘proofreading’, ‘editing capabilities’ and ‘positional information’.²⁰⁶

The DNA sequence represents, so the argument goes, something that comes from natural selection because the DNA sequence evolved for the purpose of determining the phenotype in the surviving organism.²⁰⁷ These accounts posit that the meaning or sense can be reduced to the biological function (teleosemantics).²⁰⁸ Thus, a ‘DNA molecule has a particular sequence because it specifies a particular protein ... [t]his element of intentionality comes from natural selection’.²⁰⁹ These DNA sequences are special in the sense that ‘biologists draw a distinction between two types of causal chain, genetic and environmental, or “nature” and “nurture”’.²¹⁰ In addition, ‘[f]luctuations in the environment are a source of noise in the system, not of information’ (e.g., a cake recipe will turn out slightly differently when baked in a different oven).²¹¹

DNA contains information that has been programmed by natural selection; that this information codes for the amino acid sequence of proteins; that, in a much less well understood sense, the DNA and proteins carry instructions, or a program, for the development of the organism; that natural selection of organisms alters the information in the genome; and finally, that genomic information is ‘meaningful’ in that it generates an organism able to survive in the environment in which selection has acted.²¹²

There are problems with this approach. There must be a difference between the message in the DNA sequence and the physics and chemistry of the DNA sequence. Otherwise, ‘we have chemistry or physics and not semantics’.²¹³ This difference is reputed to be ‘arbitrariness’ (or ‘gratuity’) in the sense that the form of the molecule is different to its meaning.²¹⁴ For example, Jacob and Monod demonstrated that the *lac* operon regulation was effected through inducers and repressors interacting with a regulatory protein. This changed their shape so that another part of the regulator protein bound to the DNA sequence, switching on or off *any* transcription of the DNA sequence open reading frame (the β -galactosidase).²¹⁵ This suggests that ‘there is no necessary connection between [the inducer and repressor] form (chemical composition) and meaning (genes switched on or off)’.²¹⁶ Therefore, like a symbolic language potentially conveying an indefinite number of meanings, it is ‘the symbolic nature of molecular biology that makes possible an indefinite large number of biological forms’.²¹⁷ Reduced to the idea, the function is not specifically determined just by the chemistry of the molecules.²¹⁸ This does not, however, seem credible because for DNA ‘the structure of the “message” is too closely connected to the structure of the “signal”’; therefore, the ‘arbitrariness’ is not clear,²¹⁹ and with so many causal links, the salient ones cannot be distinguished from the others, suggesting that ‘the distinction between arbitrary and non-arbitrary causal roles is just in the eye of the beholder’.²²⁰ At best, ‘arbitrariness’ is a ‘useful if elusive concept in biology’.²²¹ Another criticism is that if ‘arbitrariness’ is necessary for molecules to have meaning, then assuming DNA (and RNA) have information, they will have the relevant chemically arbitrary relations.²²² Translation may be relevantly arbitrary because any given mRNA is arbitrarily related to the proteins it specifies, but there is no arbitrariness in

²⁰⁴ Godfrey-Smith, “Misinformation.” See also Griffiths, “Lehrman’s Dictum,” 24–25.

²⁰⁵ Griffiths, “Genetic Information,” 397–398.

²⁰⁶ Rosenberg, “Darwinian Reductionism,” 99–100.

²⁰⁷ Maynard-Smith, “Concept of Information,” 189–190. See also Jablonka, “Information: Its Interpretation.”

²⁰⁸ See Kumar, “Information, Meaning, and Error in Biology,” 91–93. See also Millikan, “Language, Thought and Other Biological Categories.”

²⁰⁹ Maynard-Smith, “Concept of Information,” 189–190. See also Levy, “Information in Biology: A Fictionalist Account,” 643.

²¹⁰ Maynard-Smith, “Concept of Information,” 189.

²¹¹ Maynard-Smith, “Concept of Information,” 192.

²¹² Maynard-Smith, “Concept of Information,” 190. See also Sterelny, “The Extended Replicator.”

²¹³ Levy, “Information in Biology: A Fictionalist Account,” 643.

²¹⁴ Maynard-Smith, “Concept of Information,” 185. See also Sterelny, “The ‘Genetic Program’ Program,” 197–198.

²¹⁵ Jacobs, “Genetic Regulatory Mechanisms,” 319–328. See also Stegmann, “The Arbitrariness of the Genetic Code,” 209–214; Maynard-Smith, “Concept of Information,” 185.

²¹⁶ Maynard-Smith, “Concept of Information,” 185.

²¹⁷ Maynard-Smith, “Concept of Information,” 185. See also Stegmann, “The Arbitrariness of the Genetic Code,” 206, 218.

²¹⁸ Maynard-Smith, “Concept of Information,” 183, 185; Sterelny, “The ‘Genetic Program’ Program,” 197–198.

²¹⁹ Levy, “Information in Biology: A Fictionalist Account,” 643.

²²⁰ Godfrey-Smith, “Information, Arbitrariness, and Selection,” 203. See also Levy, “Information in Biology: A Fictionalist Account,” 643. Although Godfrey-Smith does accept that ‘[t]he idea of arbitrariness is elusive ... but it might well be useful in some contexts. And the peculiar way in which DNA has its effects on cellular processes – as much a matter of what DNA does not do, as a matter of what it does do – gives some sense to the idea that DNA is an “informational polymer”’ (p. 206).

²²¹ Godfrey-Smith, “Information, Arbitrariness, and Selection,” 205.

²²² Stegmann, “The Arbitrariness of the Genetic Code,” 218.

the ‘supposedly informational processes’ of transcribing DNA to RNA or replicating DNA to DNA.²²³ Further, if DNA sequences contain intentional information, then so must many other biological entities,²²⁴ and there is, again, parity between the different causes,²²⁵ thus:

Nucleic acid sequences and phospholipid membranes both have distinctive and essential roles in the chemistry of life and in both cases there seems no realistic substitute for them. But the facts of development do not justify assigning DNA the role of information and control while inherited membrane templates get the role of ‘material support’ for reading DNA.²²⁶

In contrast with causal information that seeks to distinguish between genetic (nature) and environmental (nurture) causes (recall the gene for achondroplasia and malnourished and underweight children),²²⁷ the intentional information seeks to distinguish between developmental genetic information and other causes.²²⁸ Preferencing causal paths between genotypes to phenotypes as different to the other causal paths between non-genetic factors (like the environment) to phenotypes²²⁹ cannot be an adequate characterisation.²³⁰ More sophisticated models positing a transmission sense of information (communications engineer’s approach)²³¹ and teleosemantic information (infotel framework)²³² also fail.²³³ The communications engineer’s approach fails because there is no account of the teleosemantic content of these signals that have been designed by natural selection.²³⁴ The infotel framework²³⁵ does, however, provide some insights, although it ultimately fails. The basic proposition is that genetic and environmental causes in development can carry inherited information because the organism has adapted a phenotype over evolutionary time, carrying information about the selection pressures in past environments. These selection pressures are ‘represented’ in the DNA sequences of the current population that correlate with those past events.²³⁵ Thus, ‘reading information carried by the genome, information that has been built up in phylogenetic time through the process of natural selection’.²³⁶ According to this model, the intentional information in the DNA sequences is the expression of the represented adaptation (the phenotype) in the current environment.²³⁷ This does not, however, adequately account for observations in practice. For example, the seed beetle *Stator limbatus* develops according to the survival rates posed by the different seeds species on which it lays eggs.²³⁸ Eggs laid on *Acacia greggii* seeds have very high rates of survival, while eggs laid on *Cercidium floridum* seeds face challenges.²³⁹ To address these challenges, the seed beetle lays fewer larger eggs on the *Cercidium floridum* seeds.²⁴⁰ Following the infotel theory, the outcome conflates teleosemantic information with a mechanistic role:

Having detected which kind of seed it is depositing eggs upon, the mother signals to the offspring to adopt one growth strategy rather than another. Using the infotel theory, we can assign the larger egg mass the indicative content ‘you are on *Cercidium floridum*’ and the imperative content ‘grow fast and get large’ ... this teleosemantic transmission information does not translate into a mechanistic explanation of development. If we ask the developmental question ‘how does the egg mass produce faster growth and larger size’ and answer ‘by transmitting to the mechanisms of development the instruction to grow fast and get large’ or ‘by transmitting to the mechanisms of development the information that the egg has been laid on *Cercidium floridum*’ it is evident how vacuous this is as an explanation.²⁴¹

²²³ Because transcribing DNA to RNA or replicating DNA to DNA are purely stereospecific covalent chemical bonding: reactions Stegmann, “The Arbitrariness of the Genetic Code,” 218.

²²⁴ See Griffiths, “Replicator II: Judgement Day”; Sterelny, “The Extended Replicator.”

²²⁵ ‘Any defensible definition of information in developmental biology is equally applicable to genetic and non-genetic causal factors in development’: Griffiths, “Genetic Information,” 396, 401, 406. See also Godfrey-Smith, “Information, Arbitrariness, and Selection,” 202; Griffiths, “Developmental Systems.”

²²⁶ Griffiths, “What is the Developmentalist Challenge?” 254. Note that DNA, RNA and proteins can be synthetically substituted through codon optimisation, nucleotide modification, epigenetic modification and protein modification: see CBD/DSI/AHTEG/2020/1/3, Annex (pp. 16–17).

²²⁷ See Maynard-Smith, “Concept of Information,” 189. See also Griffiths, “Genetic Information,” 398.

²²⁸ Griffiths, “Genetics and Philosophy,” 163.

²²⁹ Godfrey-Smith, “On the Theoretical Role of ‘Genetic Coding,’” 28.

²³⁰ See Griffiths, “Genetic Information,” 399–401 (parity), 401–403 (epigenetic inheritance), 403–406 (limited inheritance systems), 406–409 (very bad developmental switches argument). See also Kjosavik, “From Symbolism to Information.”

²³¹ Bergstrom, “The Transmission Sense of Information,” 163–171 and the references therein.

²³² Shea, “Inherited Representations,” 7–9. See also Shea, “Consumers Need Information.”

²³³ See Griffiths, “Genetics and Philosophy,” 160–174. See also Godfrey-Smith, “Senders, Receivers and Genetic Information.”

²³⁴ Griffiths, “Lehrman’s Dictum,” 26.

²³⁵ Shea, “Representation in the Genome,” 315–316, 324–328.

²³⁶ Shea, “Inherited Representations,” 3. See also Shea, “Representation in the Genome.”

²³⁷ Shea, “Representation in the Genome,” 329.

²³⁸ Griffiths, “Genetics and Philosophy,” 172 and the references therein; Griffiths, “Lehrman’s Dictum,” 28 and the references therein.

²³⁹ Griffiths, “Genetics and Philosophy,” 172; Griffiths, “Lehrman’s Dictum,” 28.

²⁴⁰ Griffiths, “Genetics and Philosophy,” 172; Griffiths, “Lehrman’s Dictum,” 28.

²⁴¹ Griffiths, “Lehrman’s Dictum,” 28. See also Griffiths, “Genetics and Philosophy,” 172.

This confounding teleosemantic information with a mechanistic role needs to clearly distinguish between developmental explanations and evolutionary explanations.²⁴² Thus, how an eye develops and how an eye works (focusing and transducing light) is different, and the DNA sequence cannot account for how an eye works as that is an evolutionary explanation for a preferred mechanism rather than a developmental mechanism.²⁴³ In the context of *Cercidium floridum*, the information in the DNA sequences is not ‘grow fast and get large’ because this intentional ‘grow fast and get large’ is neither the specified order of amino acids in the proteins nor the adaptive history of natural selection. These ‘are not mechanistic explanations of how phenotypes are constructed by the regulated expression of the genome, but evolutionary explanations of why development uses a particular mechanism to produce that outcome. They are evolutionary explanations of developmental phenotypes’.²⁴⁴

Alternatively, if the mother laying eggs on *Cercidium floridum* seeds is characterised without historical information, and the ‘grow fast and get large’ is *specifically* caused by the environmental variable and the state of the organism, then this might be adaptive information: ‘just as something needs to be adaptive in the past to be an adaptation in the future, a representation needs to have contained adaptive information in the past if it is to contain inherited information in the future’.²⁴⁵ But of course, this is really just reframing the claim for inherited intentional information as causal specificity of the DNA sequence as an actual difference-maker within the current functioning organism.

Another possibility is that DNA sequences might be considered as having a limited kind of intentional information: ‘As templates for the synthesis of macromolecules, nucleic acids determine their products in a way that is constitutive for instructions in general. It is therefore legitimate to attribute instructional content to molecular templates’.²⁴⁶ The argument here is essentially that DNA sequences are information about the linear order of the components of RNA and proteins. This might be expressed correctly or incorrectly, and the DNA sequence holds the information in the sense that the DNA sequence serves as a template for synthesising RNA and proteins.²⁴⁷ This is the same as saying the DNA sequence is Crick information with the assertion that ‘this is information in a semantic, rather than a purely correlational, sense’ because there is ‘something in the process of template-directed synthesis itself that motivates the attribution of semantic information’.²⁴⁸ This account does not progress beyond Crick information and does not support the existence of intentional information beyond the mere information of RNA and protein sequence.

DNA Sequence and Information

The adoption of information language in molecular biology²⁴⁹ coincided with the move of scientists from physics to biology²⁵⁰ and the funding from the Rockefeller Foundation (and Caltech) to promote the application of physical science to biology.²⁵¹ Perhaps importantly, physicists framed the new molecular biology in expressly informational terms, such as Erwin Schrödinger asserting in 1944 that chromosomes ‘contain in some kind of code-script the entire pattern of the individual’s future development and of its functioning in the mature state’.²⁵² This appears to have had a significant influence on the emerging molecular biology community with a ‘new vision of biology’ and providing the conceptual framework for proposing, doing and interpreting experiments, as well as a preference for determinist and informational explanations.²⁵³ The former physicist Crick’s central dogma illustrates these developments with a focus on information flows of ‘detailed, residue-by-residue, sequence information from one polymer molecule to another’.²⁵⁴ This information language has been extraordinarily successful in opening up molecular biology to reveal the molecular basis of the classical gene and the molecular networks for functioning

²⁴² “Inherited Representations,” 9, 24–27.

²⁴³ Griffiths, “Genetics and Philosophy,” 174; Griffiths, “Lehrman’s Dictum,” 29.

²⁴⁴ Griffiths, “Genetics and Philosophy,” 174.

²⁴⁵ Griffiths, “Proximate and Ultimate Information,” 90.

²⁴⁶ Stegmann, “Genetic Information as Instructional Content,” 426.

²⁴⁷ Stegmann, “Genetic Information as Instructional Content,” 429. Noting also that DNA sequence serving as a template for synthesising RNA and proteins does not accord with the observation that proteins are assisted by chaperones to correctly fold: see, for example, Balchin, “Recent Advances in Understanding.”

²⁴⁸ Stegmann, “Genetic Information as Instructional Content,” 429.

²⁴⁹ For an overview see Kay, “Who Wrote the Book of Life?,” 615–628. See also Gamow, “The Cryptographic Approach,” 63–69; Yčas, “The Protein Text,” 70–100; Augenstine, “Protein Structure and Information Content,” 103–123; Quastler, “Essays on the Use of Information Theory in Biology.”

²⁵⁰ See Morange, “History of Molecular Biology,” 67–78, 99–101. See also Fleming, “Émigré Physicists.”

²⁵¹ Kay, “The Molecular Vision of Life,” 3–57; Morange, “History of Molecular Biology,” 79–87; Kohler, “The Management of Science,” 301–304.

²⁵² Schrödinger, “What is Life?,” 21.

²⁵³ Morange, “History of Molecular Biology,” 67–78, 99–101. See also Yoxen, “Where Does Schroedinger’s ‘What is Life?’ Belong.” See also Weaver, “Molecular Biology.”

²⁵⁴ Crick, “Central Dogma of Molecular Biology,” 561.

organelles, cells and organisms. While there was hope (and still is?) that biology would be an information science,²⁵⁵ serious attempts at biosemiotics,²⁵⁶ teleosemantics²⁵⁷ and deflationary accounts²⁵⁸ have not been successful.²⁵⁹ It remains unclear whether using information language is just heuristically useful fiction,²⁶⁰ an illustrative metaphor²⁶¹ or something else. The analysis presented in this article shows that there are two related forms of information in DNA sequences, and these relate to the molecular gene rather than the classical gene:

1. Crick information—the causal specificity between DNA bases and the precise determination of the primary structure of RNA and proteins so that any cause(s) that makes a specific difference to a linear RNA or protein.
2. Other causal information—correlations between DNA bases and their downstream products as RNAs and expressed proteins and, in some cases, phenotypic traits where the causal specificity of the DNA sequence is an actual difference-maker (and this can be measured as Shannon information in bits). Importantly, the DNA sequence as an actual difference-maker needs to be determined on a case-by-case basis accounting for spatial and temporal differences.

The outcome of the large DNA sequencing projects, such as the Human Genome Project in the early 2000s,²⁶² the Haplotype Mapping (HapMap) Project,²⁶³ the 1000 Genomes Project²⁶⁴ and the *Arabidopsis* Genome Initiative,²⁶⁵ yielded very little insight into the functioning of organisms, demonstrating that the DNA sequence alone is not sufficient to improve our understanding of biology.²⁶⁶ What was required was the integration of the DNA sequence with other information about the organisation, components and processes of biological organisms and an end of the ideal of DNA sequence as the ultimate reductionist biology.²⁶⁷ The recent advances applying Shannon information theory to causal DNA sequence information have shown that the contribution of the DNA sequence needs to be assessed for every instance. Only in some cases will the DNA sequence be the most significant cause and very, very rarely the *only* cause. Recall the DSCAM coding sequence example, where the DSCAM DNA sequence is not the only causal information (and also provides no intentional information) and the specific context of the DNA sequence (*Drosophila* or human) was a critical determining factor. Put bluntly, the same DNA sequences have different amounts of information depending on their context, and this is only a limited kind of causal information.

A further limitation inherent in the DNA sequence as information is that the causal account fails to capture the directionality of information flows from DNA sequences (genotypes) to phenotypes.²⁶⁸ Recall the central dogma that ‘once “information” has passed into protein *it cannot get out again*’ (emphasis in original).²⁶⁹ The Shannon information theory, however, would posit that ‘the amount of information that knowing the genotype *G* provides about the phenotype *P* is always exactly equal to the amount of information that knowing the phenotype *P* provides about the genotype *G*’.²⁷⁰ However, for DNA sequences, this is a privileged directional flow of information from DNA sequence to protein where the DNA sequence informs the protein but

²⁵⁵ George Williams in 1992 posited a ‘codical domain’ for biology on the basis that ‘DNA is the medium, not the message. A gene is not a DNA molecule; it is the transcribable information coded by the molecule ... the gene is a packet of information, not an object’: Sterelny and Griffiths, “Sex and Death,” 100 citing Williams, “Plan and Purpose in Nature.”

²⁵⁶ See, for example, Hoffmeyer, “Biosemiotics”; Artmann, “Computing Codes,” 209–234; Witzany, “Life: The Communicative Structure.”

²⁵⁷ See, for example, Shea, “Representation in the Genome”; Maynard-Smith, “Concept of Information.”

²⁵⁸ See, for example, Sarkar, “Genes Encode Information,” 259–262. See also Stegmann, “DNA, Inference and Information.”

²⁵⁹ There are also others that remain incomplete: see, for example, Kjosavik, “Genes, Structuring Powers and the Flow”; Bogen, “Mechanistic Information.”

²⁶⁰ “Information-talk is serious but it isn’t literally true”: Levy, “Information in Biology: A Fictionalist Account,” 647.

²⁶¹ ‘The present atmosphere information talk is only applied to genes, makes that way of talk mislead.’ Griffiths, “Genetic Information,” 410.

²⁶² See Venter, “Sequence of the Human Genome”; Human Genome Sequencing Consortium, “Initial Sequencing and Analysis of the Human Genome.”

²⁶³ Gibbs, “International HapMap Project.”

²⁶⁴ 1000 Genomes Project Consortium, “A Map of Human Genome Variation.”

²⁶⁵ Arabidopsis Genome Initiative, “Analysis of the Genome Sequence.”

²⁶⁶ See, for example, Gibbs, “Human Genome Project Changed Everything,” 576.

²⁶⁷ ‘The Gene Ontology (GO) project provides an example of evidence-supported annotations to describe the biological roles of individual genomic products (like genes, proteins, ncRNAs, complexes) by classifying them using our ontologies’: Gene Ontology Consortium, “Gene Ontology Consortium.”

²⁶⁸ Griffiths, “Genetic Information,” 398; Godfrey-Smith, “On the Theoretical Role,” 41–42. See also Bergstrom and Rosvall, “The Transmission Sense of Information,” 161–162.

²⁶⁹ Crick, “On Protein Synthesis,” 153. See also Crick, “Central Dogma of Molecular Biology,” 562. See also Sarabhai, “Co-linearity of the Gene.”

²⁷⁰ Bergstrom, “The Transmission Sense of Information,” 161, 164–165.

not vice-versa.²⁷¹ There is a redundancy in the code so that from a protein, there are a range of possible RNAs and, consequently, DNA sequences. While this does not invalidate the application of Shannon information theory to causal information in DNA sequences, it suggests some caution in applying this abstracted theory to material DNA sequences.

This analysis shows that causal and intentional information is most often a metaphor when discussing DNA sequences, except for the limited and special circumstances of Crick information and where the DNA sequence is the causal actual difference-maker. There is no doubt that information metaphors are useful, and they have expounded the molecular gene in both the scientific and popular imagination. These metaphors may, however, have a more subtle role that has framed the way many discussions around DNA sequence information are limited in the CBD and Nagoya Protocol forums, conflating the ideals of classical genes and molecular genes:

The very structure of a typical genetics education endows the character-makers conception with independent life, by investing it with heuristic power. Begin your education in genetics with Mendel's peas, and you will learn not merely about a case where, you are told, binary characters are determined by genes for those characters, and by nothing else. You will learn too that many apparently more complicated cases can be made tractable by treating them in the first instance like Mendel's peas. (And if you don't learn that, you won't pass.) Of course, you will go on to learn about all sorts of exceptions to your rule of thumb, and the reasons why those exceptions are the way they are: the effects of other genes, epigenetic modifications, the interplay of development and environment, chance. By the end of your education, you will know, of course, that 'it's not all in the genes', and become annoyed with anyone who suggests that you think otherwise. But the Mendelian, treat-'em-like-the-peas rule of thumb will remain in place. It will guide your reasoning and even – in the way of heuristics – perhaps your unreasoned reflections and reactions too, with much reinforcement from the wider culture in the form of gene-personifying 23-and-Me ads, 'gene for' discovery stories, jargon talk of what is in an organization's DNA, and so on. You will affirm genes-for-characters determinism in your actions and attitudes while rejecting it if asked about it, because you know that it's false (footnotes excluded).²⁷²

That this is problematic is perhaps best illustrated by a limited ABS scheme directed to human influenza viruses. The World Health Organisation of the United Nations' *Pandemic Influenza Preparedness Framework* (PIP Framework)²⁷³ defines 'gene sequences' as 'the order of nucleotides found in a molecule of DNA or RNA. They contain the genetic information that determines the biological characteristics of an organism or a virus'.²⁷⁴ As the analysis in this article clearly demonstrated, 'the order of nucleotides found in a molecule of DNA or RNA' does not necessarily 'contain the genetic information that determines the biological characteristics of an organism or a virus'. Where they do 'contain ... genetic information', this is limited to the Crick information and other causal information that is an actual difference-maker that needs to be determined on a case-by-case basis accounting for spatial and temporal differences. Clearly distinguishing between the ideals of classical and molecular genes is critical to the DSI discussions so that simplistic and ultimately misleading conceptions of DNA sequence do not undermine the role and place of information *in* DNA sequences. The next question, with these clear limits on thinking about information *in* DNA sequences, is how these matters should be addressed in the context of the DSI debates at the CBD and Nagoya Protocol forums.

Regulating DSI

Information under existing CBD and Nagoya Protocol ABS obligations, however described and defined, has been addressed through:

1. The ABS contract as a term and condition of prior informed consent and/or mutually agreed terms.²⁷⁵ In addition to the general obligations set out in the CBD and Nagoya Protocol, there are general trade standards established by the World Trade Organisation's *Agreement on Trade Related Aspects of Intellectual Property Rights* (TRIPS) that will be implemented through these ABS contracts as terms and conditions of the agreement. This includes laws addressing

²⁷¹ Godfrey-Smith, "Information in Biology," 106. For an alternative see Yockey, "An Application of Information Theory." Note also that multiple different DNA sequences can encode the same protein because of the degeneracy of the genetic code (surjective).

²⁷² Radick, "Making Sense of Mendelian Genes," 308.

²⁷³ WHA64.5 (PIP Framework).

²⁷⁴ PIP Framework, Art. 4.2.

²⁷⁵ CBD, Arts. 15.4 and 15.5; Nagoya Protocol, Arts. 5 and 6.

copyrights,²⁷⁶ patents,²⁷⁷ secrets,²⁷⁸ confidential information²⁷⁹ and information submitted for governmental approval.²⁸⁰

2. The general obligation promoting the exchange of information,²⁸¹ the CBD Clearing House Mechanism²⁸² and the Nagoya Protocol Access and Benefit Sharing Clearing.²⁸³

The challenge for addressing DSI is whether these existing arrangements are suitable or whether further or different measures are necessary. This is the important distinction because the general information obligations promote the disclosure and exchange of information, while proposals to address DSI treat DSI as a derivative of the materials within the ABS transaction itself, which becomes a distinct commodity with a value that an ABS scheme attempts to translate into definable benefits.²⁸⁴ The mischief that needs to be addressed here is using DSI without the physical genetic materials, potentially avoiding the ABS prior informed consent and/or mutually agreed terms requirements, including the benefit-sharing.²⁸⁵ The preferable outcome is a simple, efficient and effective multilateral agreement balancing access and benefit-sharing that delivers fair and appropriate benefits from the access and utilization of DSI (whatever that might be).²⁸⁶

The analysis in this article shows that DNA sequences *only* have Crick information and *may* have causative information that would need to be assessed for each and every sequence. Put simply, there is limited information *in* DNA sequences, and consequently, there is lots of information *about* DNA sequences. This distinction and the limitations of treating DNA sequences as information *per se* become readily apparent in applying the DSI groupings proposed by the AHTEG-DSI commissioned study on the concept and scope of DSI and how DSI was currently used.²⁸⁷ This study's groupings were, noting that these are expressed in the sense of the molecular gene and not the classical gene ideal:²⁸⁸

1. 'Group 1 – Narrow: DNA and RNA'—'a narrow scope or proximity to the genetic resource and is limited to nucleotide sequence information associated with transcription'.²⁸⁹ This is Crick information and limited to the extent that the DNA sequence information is the precise determination of base sequence in the nucleic acids (DNA and RNA) but less than the full Crick information because it does not include the precise determination of amino acid residues sequence in the protein.
2. 'Group 2 – Intermediate: (DNA and RNA) + proteins'—'an intermediate scope and extends to protein sequences, thus comprising information associated with transcription and translation'²⁹⁰ (but only in one direction DNA to protein and less precision or confidence for protein to DNA because of the redundant genetic code). This is Crick information as the DNA sequence information is the precise determination of base and amino acid residues sequence in the nucleic acids (DNA and RNA) and protein. This will also include information *about* the DNA sequence that is correlated with the DNA sequence but is not an actual difference-maker (recall the DSCAM example above).
3. 'Group 3 – Intermediate: (DNA, RNA and proteins) + metabolites'—'a wider intermediate scope or proximity to the genetic resource and extends to metabolites and biochemical pathways, thus comprising information associated with transcription, translation and biosynthesis'.²⁹¹ This is Crick information and additional causal information where the causal specificity of the DNA sequence is an actual difference-maker for the metabolites. A DNA sequence as an actual difference-maker, however, would need to be assessed on a case-by-case basis accounting for spatial and

²⁷⁶ See TRIPS, Arts. 9–14.

²⁷⁷ See TRIPS, Arts. 27–34.

²⁷⁸ See TRIPS, Art. 39.2.

²⁷⁹ See TRIPS, Art. 39.1.

²⁸⁰ See TRIPS, Art. 39.3.

²⁸¹ CBD, Arts. 16.1 and 17.1.

²⁸² CBD, Art. 18.3.

²⁸³ Nagoya Protocol, Art. 14.

²⁸⁴ Lawson, "The Future of Information," 104.

²⁸⁵ See Lawson, "Information as the Latest Site," 19–26.

²⁸⁶ This rejects the proposed "Option 5: No Benefit Sharing from DSI": CBD/WG2020/3/4, Annex II (p. 17).

²⁸⁷ CBD/DSI/AHTEG/2020/1/3.

²⁸⁸ CBD/DSI/AHTEG/2020/1/3, Annex (p. 32).

²⁸⁹ CBD/DSI/AHTEG/2020/1/3, [12] and Annex (pp. 32 and 38).

²⁹⁰ CBD/DSI/AHTEG/2020/1/3, [12] and Annex (p. 32).

²⁹¹ CBD/DSI/AHTEG/2020/1/3, [12] and Annex (pp. 32 and 44).

temporal differences. Again, this will also include information *about* the DNA sequence that is correlated with the DNA sequence but is not an actual difference-maker.

4. ‘Group 4 – Broad: (DNA, RNA, protein, metabolites) + traditional knowledge, ecological interactions, [and so on]’— ‘the broadest scope or weakest proximity to the underlying genetic resource and extends to behavioural data, information on ecological relationships and traditional knowledge, thus comprising information associated with transcription, translation and biosynthesis, as well as downstream subsidiary information concerning interactions with other genetic resources and the environment as well as its utilization, among other subsidiary information’.²⁹² This is Crick information and additional causal information where the causal specificity of the DNA sequence is an actual difference-maker for the metabolites. Again, a DNA sequence as an actual difference-maker would need to be assessed on a case-by-case basis accounting for spatial and temporal differences. This grouping will also include information *about* the DNA sequence, such as traditional knowledge and ecological interactions, that is correlated with the DNA sequence but is not an actual difference-maker.

The analysis clearly shows that if the DNA sequence is treated as a derivative of the materials within the ABS transaction itself and becomes a distinct commodity with a value that the ABS scheme attempts to translate into definable benefits (as some countries have already done through laws²⁹³ or as mandatory terms and conditions as part of prior informed consent and/or mutually agreed terms addressing DSI),²⁹⁴ privileging the DNA sequence will undervalue the other non-DNA sequence contributions to phenotypes. This will potentially limit the kinds and values of research and development on the other causal contributions.²⁹⁵ Put simply, and demonstrated by the AHTEG-DSI commissioned study, the common understanding of genetics privileging bottom-up information flowing from DNA sequences is pervasive but a misleading base to found a legislative, administrative and policy ABS scheme. This will inevitably undermine the purpose of ABS schemes to deliver benefits for conservation and sustainable uses and the integrity of ABS schemes because the value of the DNA sequence that is the actual difference-maker is not being assessed (including accounting for spatial and temporal differences), and the value of most other DNA sequences is being overvalued and tied up with complex law, policy and processes. Predictably, this leads to perverse outcomes by controlling the potential uses of information or reducing the incentives to use information in new and innovative ways, and consequently for the conservation and sustainable use of biodiversity.²⁹⁶ Further, merely leaving DNA sequences as DSI to be resolved within the current ABS scheme is problematic because this will create a complicated matrix of different laws, policies and practices as each CBD Contracting Party and Nagoya Protocol Party implements their own approaches, perpetuating these perverse outcomes.²⁹⁷ This suggests that DSI is not a suitable target for regulation because it both simplistically privileges the bottom-up information flowing from DNA sequences and presumes all DNA sequences and their extensions (e.g., information associated with transcription, translation and biosynthesis) are sufficiently valuable to warrant complex ABS negotiation and agreement-making.

The dilemma for the CBD Contracting States and Nagoya Protocol Parties is whether to persevere in crafting the regulation of DNA sequence information to the kinds of nuances identified in the presented analysis about Crick information and causal actual difference-maker information and the inherent problems of deciding thresholds for how much actual difference-making has value. This might be possible but will likely directly conflict with the ideal of a simple, efficient and effective multilateral agreement because of the difficult assessments about whether the information is actually valuable²⁹⁸ and how and when to impose those obligations (e.g., agreed standard material transfer terms and conditions) and deliver the benefits.²⁹⁹ The likely most simple, efficient and effective compromise is to accept that a more general multilateral solution will have some inefficiencies capturing other kinds of information (e.g., correlated but not actual difference-maker information) but that *also* delivers certainty and predictability to those accessing DSI, externalises the value of benefit-sharing from the complexities of a matrix of different laws, policies and practices among the CBD Contracting States and Nagoya Protocol Parties and facilitates

²⁹² CBD/DSI/AHTEG/2020/1/3, { 12} and Annex (p. 32).

²⁹³ See CBD/DSI/AHTEG/2020/1/5, 10–11.

²⁹⁴ See, for example, CBD/DSI/AHTEG/2020/1/5, 21.

²⁹⁵ This rejects the proposed ‘Option 1: DSI fully integrated into the Convention on Biological Diversity and the Nagoya Protocol’: CBD/WG2020/3/4, Annex II (p. 15).

²⁹⁶ This rejects the proposed ‘Option 0: Status Quo’: CBD/WG2020/3/4, Annex II (p. 15). See also Lawson, “Patent Laws Will Undermine Access,” 401.

²⁹⁷ This also rejects the proposed ‘Option 0: Status Quo’: CBD/WG2020/3/4, Annex II (p. 15). See also Sara, “A Need for Recalibration,” 2; Mekonnen and Spielman, “Changing Patterns in Genebank Acquisitions,” 4–9.

²⁹⁸ See Lawson, “Information as the Latest Site,” 13–14.

²⁹⁹ This rejects the proposed ‘Option 2: Standard [Material Transfer Agreement]’: CBD/WG2020/3/4, Annex II (p. 15). See also Sara, “A Need for Recalibration,” 2.

easy access to DNA sequences for *any* uses.³⁰⁰ If the solution is to externalise the value of benefits in a multilateral agreement that is made separately from the ABS transaction (decoupling), this might include either a payment (e.g., a charge, levy or tax)³⁰¹ or other non-monetary benefits (e.g., research collaborations, training, knowledge platforms, technology transfer and technology co-development)³⁰² or a combination of these measures. The guiding principles for such a multilateral agreement are that users have certainty and predictability about their obligations (clear provenance) and facilitated, easy access to DNA sequences and are aware that regulatory options need to be considered and assessed in the context of *both* the classical gene and the molecular gene.

Conclusions

The starting point for working out whether there is information *in* DNA sequences was to distinguish between the classical gene and the molecular gene and appreciate the success of the reductionist, bottom-up account based in a physical sciences methodology that *is* DNA as the molecular gene, so the genotypes are the *causes* of phenotypes. Recall, however, that while classical genes and molecular genes might not be separate theories, conflating the two levels of resolution privileging the molecular gene overlooks the other factors (including non-genetic factors) that affect genotypes. This becomes important in appreciating the distinctions about the kinds of information *in* DNA sequences. There is no intentional or semantic information in DNA sequences. There is *only* causative information—Crick information in the DNA sequence specifying the linear order of RNA and proteins and other causal information where that DNA sequence makes an actual causal difference to the observed phenotype. These actual causal differences to the observed phenotype also apply to *all* other biological systems having an effect so that the quantum of the DNA sequence causation will vary from nothing to a lot depending on the particular spatial and temporal circumstances (recalling the DSCAM example). Privileging the DNA sequence undervalues all these other causes, limits the kinds and values of research and development about the other causal contributions and will likely reduce research and development using DNA sequences subject to ABS measures. The predictable consequence is to undermine the CBD's objectives of conservation and sustainable use of biodiversity. Perhaps most importantly, the analysis here cautions against both the 'genes-for-characters determinism in ... actions and attitudes'³⁰³ for those proposing DSI regulation and highlights the potential perverse outcomes from privileging the DNA sequence as a molecular gene rather than the more complicated classical gene. The analysis here again shows that it is time to rethink ABS and that there needs to be new approaches to make scientific research and commercialisation more equitable and sustainable.³⁰⁴

Acknowledgements

The author thanks Fran Humphries, Michelle Rourke, Marcel Jaspars and the Journal's reviewers and editors for their insightful comments and suggestions.

Bibliography

- 1000 Genomes Project Consortium. "A Map of Human Genome Variation from Population-scale Sequencing." *Nature* 467, no 7319 (2010): 1061–1073. <https://www.nature.com/articles/nature09534>
- Adams, Mark, Susan Celniker, Robert Holt, Cheryl Evans, Jeannine Gocayne, Peter Amanatides, Steven Scherer et al. "The Genome Sequence of *Drosophila melanogaster*." *Science* 287, no 5461 (2000): 2185–2195. <https://doi.org/10.1126/science.287.5461.2185>
- Allen, Garland. "Naturalists and Experimentalists: The Genotype and The Phenotype." *Studies in the History of Biology* 3 (1979): 179–209.

³⁰⁰ This accepts the proposed 'Option 3: No [Prior Informed Consent], No [Material Transfer Agreement]' and 'Option 4: Enhanced technical and scientific cooperation': CBD/WG2020/3/4, Annex II (pp. 16–17).

³⁰¹ This is the proposed 'Option 3: No [Prior Informed Consent], No [Material Transfer Agreement]': CBD/WG2020/3/4, Annex II (pp. 16–17).

³⁰² This is the proposed 'Option 4: Enhanced technical and scientific cooperation': CBD/WG2020/3/4, Annex II (p. 17).

³⁰³ Radick, "Making Sense of Mendelian Genes," 308.

³⁰⁴ Laird, "Rethink the Expansion of ABS," 1202.

- Anfinsen, Christian, Elie Haber, Michael Sela and Frederick White. "The Kinetics of Formation of Native Ribonuclease During Oxidation of the Reduced Polypeptide Chain." *Proceedings of the National Academy of Sciences* 47, no 9 (1961): 1309–1314. <https://dx.doi.org/10.1073%2Fpnas.47.9.1309>
- Arabidopsis Genome Initiative. "Analysis of the Genome Sequence of the Flowering Plant *Arabidopsis thaliana*." *Nature* 408, no 6814 (2000): 796–815. <https://doi.org/10.1038/35048692>
- Artmann, Stefan. "Computing Codes Versus Interpreting Life." In *Introduction to Biosemiotics*, edited by Marcello Barbieri, 209–234. Dordrecht: Springer, 2008.
- Augenstine, L. "Protein Structure and Information Content." In *Symposium on Information Theory in Biology*, edited by Hubert Yockey, 103–123. New York: Pergamon Press, 1956.
- Avery, Oswald, Colin MacLeod and Maclyn McCarty. "Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types: Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type III." *Journal of Experimental Medicine* 79, no 2 (1944): 137–158. <https://doi.org/10.1084/jem.79.2.137>
- Baetu, Tudor. "A Defense of Syntax-Based Gene Concepts in Postgenomics: Genes as Modular Subroutines in the Master Genomic Program." *Philosophy of Science* 78, no 5 (2012): 712–723. <https://doi.org/10.1086/662261>
- Balchin, David, Manajit Hayer-Hartl and Ulrich Hartl. "Recent Advances in Understanding Catalysis of Protein Folding by Molecular Chaperones." *FEBS Letters* 594, no 17 (2020): 2770–2781. <https://doi.org/10.1002/1873-3468.13844>
- Baltimore, David. "RNA-dependent DNA Polymerase in Virions of RNA Tumour Viruses." *Nature* 226, no 5252 (1970): 1209–1211. <https://doi.org/10.1038/2261209a0>
- Barbieri, Marcello. "The Definitions of Information and Meaning: Two Possible Boundaries between Physics and Biology." *Biology Forum* 97 (2004): 91–110.
- Barbieri, Marcello. "Life and Semiosis: The Real Nature of Information and Meaning." *Semiotica* 158 (2006): 233–254. <https://doi.org/10.1515/SEM.2006.007>
- Beadle, George and Edward Tatum. "Genetic Control of Biochemical Reactions in Neurospora." *Proceedings of the National Academy of Sciences* 27, no 11 (1941): 499–506. <https://doi.org/10.1073/pnas.27.11.499>
- Ben-Naim, Arieh. "Entropy and Information Theory: Uses and Misuses." *Entropy* 21, no 12 (2019): 1170. <https://doi.org/10.3390/e21121170>
- Benzer, Seymour. "Fine Structure of a Genetic Region in Bacteriophage." *Proceedings of the National Academy of Science* 41, no 6 (1955): 344–354. <https://dx.doi.org/10.1073%2Fpnas.41.6.344>
- Bergstrom, Carl and Martin Rosvall. "The Transmission Sense of Information." *Biology and Philosophy* 26, no 2 (2009): 159–176. <https://doi.org/10.1007/s10539-009-9180-z>
- Bockmann, Flávio Alicino, Miguel Trefaut Rodrigues, Tiana Kohlsdorf and Lorian Cobra Straker. "Brazil's Government Attacks Biodiversity." *Science* 360, no 3691 (2018): 865. <https://doi.org/10.1126/science.aat7540>
- Bogen, Jim and Peter Machamer. "Mechanistic Information and Causal Continuity." In *Causality in the Sciences*, edited by Phyllis McKay Illari, Federica Russo and Jon Williamson, 845–864. Oxford: Oxford University Press, 2011.
- Boveri, Theodor. *Über Mehrpolige Mitosen als Mittel zur Analyse des Zellkerns*. Würzburg: A. Stuber, 1902.
- Bowler, Peter. *Evolution: The History of an Idea*. 3rd edition. San Francisco: University of California Press, 2003.
- Calcott, Brett, Arnaud Pocheville and Paul Griffiths. "Signals That Make a Difference." *British Journal for the Philosophy of Science* 71, no 1 (2020): 233–258. <https://doi.org/10.1093/bjps/axx022>
- Camacho, Polo. "The Central Dogma is Empirically Inadequate ... No Matter How We Slice It." *Philosophy, Theory and Practice in Biology* 11 (2019): 6. <https://doi.org/10.3998/ptpbio.16039257.0011.006>
- Carlson, Elof. *The Gene: A Critical History*. Philadelphia: Saunders, 1966.
- Carlson, Elof. *Mendel's Legacy: The Origin of Classical Genetics*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 2004.
- Cavalli, Giacomo and Edith Heard. "Advances in Epigenetics Link Genetics to the Environment and Disease." *Nature* 571, no 7766 (2019): 489–499.
- C. elegans* Sequencing Consortium. "Genome Sequence of the Nematode *C. elegans*: A Platform for Investigating Biology." *Science* 282, no 5396 (1998): 2012–2018. <https://doi.org/10.1126/science.282.5396.2012>
- Chargaff, Erwin. "Chemical Specificity of Nucleic Acids and Mechanism of their Enzymatic Degradation." *Experientia* 6, no 6 (1950): 201–209. <https://doi.org/10.1007/bf02173653>
- Cohen, Stanley, Annie Chang, Herbert Boyer and Robert Helling. "Construction of Biologically Functional Bacterial Plasmids *In Vitro*." *Proceedings of the National Academy of Sciences* 70, no 11 (1973): 3240–3244. <https://doi.org/10.1073/pnas.70.11.3240>
- Colucci-D'amato, Luca, Vincenzo Bonavita and Umberto di Porzio. "The End of the Central Dogma of Neurobiology: Stem Cells and Neurogenesis in Adult CNS." *Neurological Sciences* 27, no 4 (2006): 266–270. <https://doi.org/10.1007/s10072-006-0682-z>

- Cooper, Stephen. "The Central Dogma of Cell Biology." *Cell Biology International Reports* 5, no 6 (1981): 539–549. [https://doi.org/10.1016/S0309-1651\(81\)80002-1](https://doi.org/10.1016/S0309-1651(81)80002-1)
- Crick, Francis. "On Protein Synthesis." *Symposia of the Society for Experimental Biology* 12 (1958): 138–163.
- Crick, Francis, Leslie Barnett, Sydney Brenner and Richard Watts-Tobin. "General Nature of the Genetic Code for Proteins." *Nature* 192, no 4809 (1961): 1227–1232.
- Crick, Francis. "The Genetic Code – Yesterday, Today and Tomorrow." *Cold Spring Harbor Symposium Quarterly Biology* 31 (1966): 3–9.
- Crick, Francis. "Central Dogma of Molecular Biology." *Nature* 277, no 5258 (1970): 561–563.
- Darwin, Charles. *On the Origin of Species*. 2nd edition. London: John Murray, 1860.
- Dretske, Fred. "Epistemology and Information." In *Handbook of the Philosophy of Science*, Volume 8, edited by Pieter Adriaans and Johan van Benthem, 29–47. Amsterdam: Elsevier, 2008.
- Edwards, Anthony. "Robert Heath Lock and His Textbook of Genetics, 1906." *Genetics* 194, no 3 (2013): 529–537. <https://doi.org/10.1534/genetics.113.151266>
- Fabris, Francesco. "Shannon Information Theory and Molecular Biology." *Journal of Interdisciplinary Mathematics* 12, no 1 (2009): 41–87. <https://doi.org/10.1080/09720502.2009.10700611>
- Falk, Raphael. "Mendel's Impact." *Science in Context* 19, no 2 (2006): 215–236. <https://doi.org/10.1017/S026988970600086X>
- Falk, Raphael. *Genetic Analysis: A History of Genetic Thinking*. Cambridge: Cambridge University Press, 2009.
- Fantini, Bernardino. "Of Arrows and Flows: Causality, Determination and Specificity in the Central Dogma of Molecular Biology." *History and Philosophy of the Life Sciences* 28, no 4 (2006): 567–593.
- Fleischmann, Robert, Mark Adams, Owen White, Rebecca Clayton, Ewen Kirkness, Anthony Kerlavage, Carol Bult et al. "Whole-Genome Random Sequencing and Assembly of *Haemophilus influenzae* Rd." *Science* 269, no 5223 (1995): 496–512. <https://doi.org/10.1126/science.7542800>
- Fleming, Donald. "Émigré Physicists and the Biological Revolution." *Perspectives in American History* 2 (1968): 176–213.
- Fogle, Thomas. "Information Metaphors and the Human Genome Project." *Perspectives in Biology and Medicine* 38, no 4 (1995): 535–547. <https://doi.org/10.1353/pbm.1995.0060>
- Fox Keller, Evelyn. *Refiguring Life: Metaphors of Twentieth Century Biology*. New York: Columbia University Press, 1995.
- Fuerst, John. "The Role of Reductionism in the Development of Molecular Biology: Peripheral or Central?" *Social Studies of Science* 12, no 2 (1982): 241–278. <https://doi.org/10.1177%2F030631282012002003>
- Gaffney, Jim, Redeat Tibebe, Rebecca Bart, Getu Beyene, Dejene Girma, Ndjiod ArdoKae, Emma S. Mace et al. "Open Access to Genetic Sequence Data Maximizes Value to Scientists, Farmers and Society." *Global Food Security* 26 (2020): 100411. <https://doi.org/10.1016/j.gfs.2020.100411>
- Gamow, George and Martynas Yčas. "The Cryptographic Approach to the Problem of Protein Synthesis." In *Symposium on Information Theory in Biology*, edited by Hubert Yockey, 63–69. New York: Pergamon Press, 1956.
- Gene Ontology Consortium. "Gene Ontology Consortium: Going Forward." *Nucleic Acids Research* 43, no D1 (2015): D1049–D1056. <https://doi.org/10.1093/nar/gku1179>
- Gibbs, Richard. "The Human Genome Project Changed Everything." *Nature Reviews Genetics* 21, no 10 (2020): 575–576. <https://doi.org/10.1038/s41576-020-0275-3>
- Gibbs, Richard, John Belmont, Paul Hardenbol, Thomas Willis and Fuli Yu. "The International HapMap Project." *Nature* 426, no 6968 (2003): 789–796.
- Godfrey-Smith, Peter. "Misinformation." *Canadian Journal of Philosophy* 19, no 4 (1989): 533–550. <https://doi.org/10.1080/00455091.1989.10716781>
- Godfrey-Smith, Peter. "Genes and Codes: Lessons from the Philosophy of Mind?" In *Biology Meets Psychology: Constraints, Conjectures, Connections*, edited by Valerie Hardcastle, 305–331. Cambridge: MIT Press, 1999.
- Godfrey-Smith, Peter. "Information, Arbitrariness, and Selection: Comments on Maynard Smith." *Philosophy of Science* 67, no 2 (2000): 202–207. <https://doi.org/10.1086/392770>
- Godfrey-Smith, Peter. "On the Theoretical Role of 'Genetic Coding.'" *Philosophy of Science* 67, no 1 (2000): 26–44. <https://doi.org/10.1086/392760>
- Godfrey-Smith, Peter. "Genes Do Not Encode Information for Phenotypic Traits." In *Contemporary Debates in Philosophy of Science*, edited by Christopher Hitchcock, 275–289. Oxford: Blackwell Publishing, 2003.
- Godfrey-Smith, Peter. "Information in Biology." In *The Cambridge Companion to the Philosophy of Biology*, edited by David Hull and Michael Ruse, 103–119. Cambridge: Cambridge University Press, 2008.
- Godfrey-Smith, Peter. "Senders, Receivers and Genetic Information: Comments on Bergstrom and Rosvall." *Biology and Philosophy* 26, no 2 (2011): 177–181.
- Goffeau, A., B. Barrell, H. Bussey, R. Davis, B. Dujon, H. Feldman et al. "Life with 6000 Genes." *Science* 274, no 5287 (1996): 546–567. <https://doi.org/10.1126/science.274.5287.546>
- Grice, H. Paul. "Meaning." *Philosophical Review* 66, no 3 (1957): 377–388.

- Griffiths, Paul. "Genetic Information: A Metaphor in Search of a Theory." *Philosophy of Science* 68, no 3 (2001): 394–412. <https://doi.org/10.1086/392891>
- Griffiths, Paul. "The Fearless Vampire Conservator: Philip Kitcher, Genetic Determinism, and the Informational Gene." In *Genes in Development: Rethinking the Molecular Paradigm*, edited by Eva Neumann-Held and Christoph Rehmann-Sutter, 175–198. Durham, NC: Duke University Press, 2006.
- Griffiths, Paul. "Lehrman's Dictum: Information and Explanation in Developmental Biology." *Developmental Psychobiology* 55, no 1 (2013): 22–32. <https://doi.org/10.1002/dev.21087>
- Griffiths, Paul. "Proximate and Ultimate Information in Biology." In *The Philosophy of Philip Kitcher*, edited by Mark Couch and Jessica Pfeifer, 74–91. Oxford: Oxford University Press, 2016.
- Griffiths, Paul. "Genetic, Epigenetic and Exogenetic Information in Development and Evolution." *Interface Focus* 7 (2017): 20160152. <https://doi.org/10.1098/rsfs.2016.0152>
- Griffiths, Paul and Russell Gray. "Developmental Systems and Evolutionary Explanation." *Journal of Philosophy* 91, no 6 (1994): 277–304.
- Griffiths, Paul and Russell Gray. "Replicator II: Judgement Day." *Biology and Philosophy* 12, no 4 (1997): 471–492.
- Griffiths, Paul and Russell Gray. "Three Ways to Misunderstand Developmental Systems Theory." *Biology and Philosophy* 20, no 2–3 (2005): 417–425.
- Griffiths, Paul and Robin Knight. "What is the Developmentalist Challenge?" *Philosophy of Science* 62, no 2 (1998): 253–258. <https://doi.org/10.1086/392636>
- Griffiths, Paul and Karola Stotz. *Genetics and Philosophy: An Introduction*. Cambridge: Cambridge University Press, 2013.
- Griffiths, Paul, Arnaud Pocheville, Brett Calcott, Karola Stotz, Hyunju Kim and Rob Knight. "Measuring Causal Specificity." *Philosophy of Science* 82, no 4 (2015): 529–555. <https://doi.org/10.1086/682914>
- Hamburger, Viktor. "Wilhelm Roux: Visionary with a Blind Spot." *Journal of the History of Biology* 30, no 2 (1997): 229–238. <https://doi.org/10.1023/a:1004231618837>
- Hammond, Edward and Lim Li Ching. "Discussions on Sequence Information Unravel at Biodiversity Meet, Third World Network." 19 July 2018. <http://www.twn.my/title2/biotk/2018/btk180702.htm>
- Hershey, Alfred and Martha Chase. "Independent Functions of Viral Protein and Nucleic Acid in Growth of Bacteriophage." *Journal of General Physiology* 36, no 1 (1952): 39–56. <https://doi.org/10.1085/jgp.36.1.39>
- Hoffmeyer, Jesper. *Biosemiotics: An Examination into the Signs of Life and the Life of Signs*. Chicago: University of Scranton Press, 2008.
- Hofmeister, Wilhelm. "Ueber die Entwicklung des Pollens." *Botanische Zeitung* 6, no 23 (1848): 425–434.
- Humphries, Fran, Clare Morrison, Charles Lawson, Anastasia Kolesnikova, Sarah Laird and Rachel Wynberg. *Survey of Access and Benefit-Sharing Country Measures Accommodating the Distinctive Features of Genetic Resources for Food and Agriculture and Associated Traditional Knowledge*. Rome: FAO, 2021.
- Humphries, Fran, Hiroko Muraki Gottlieb, Sarah Laird, Rachel Wynberg, Charles Lawson, Michelle Rourke et al. "A Tiered Approach to the Marine Genetic Resource Governance Framework under the Proposed UNCLOS Agreement for Biodiversity Beyond National Jurisdiction." *Marine Policy* 122 (2020): 103910.
- International Human Genome Sequencing Consortium. "Initial Sequencing and Analysis of the Human Genome." *Nature* 409, no 6822 (2001): 860–921.
- Jablonka, Eva. "Information: Its Interpretation, Its Inheritance and Its Sharing." *Philosophy of Science* 69, no 4 (2002): 578–605. <https://doi.org/10.1086/344621>
- Jackson, David, Robert Symons and Paul Berg. "Biochemical Method for Inserting New Genetic Information into DNA of Simian Virus 40: Circular SV40 DNA Molecules Containing Lambda Phage Genes and the Galactose Operon of *Escherichia coli*." *Proceedings of the National Academy of Sciences* 69, no 10 (1972): 2904–2909. <https://doi.org/10.1073/pnas.69.10.2904>
- Jacobs, François and Jacques Monod. "Genetic Regulatory Mechanisms in the Synthesis of Proteins." *Journal of Molecular Biology* 3, no 3 (1961): 318–356. [https://doi.org/10.1016/S0022-2836\(61\)80072-7](https://doi.org/10.1016/S0022-2836(61)80072-7)
- Johannsen, Wilhelm. "The Genotype Conception of Heredity." *American Naturalist* 45, no 531 (1911): 129–159. <https://doi.org/10.1086/279202>
- Judson, Horace. *The Eighth Day of Creation: The Makers of the Revolution in Biology*. New York: Simon and Schuster, 1979.
- Jumper, John, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger et al. "Highly Accurate Protein Structure Prediction with AlphaFold." *Nature* 596, no 7873 (2021): 583–589.
- Kalmus, H. "A Cybernetical Aspect of Genetics." *Journal of Heredity* 41 (1950): 19–22. <https://doi.org/10.1093/oxfordjournals.jhered.a106044>
- Kaplan, Donald and Todd Cooke. "The Genius of Wilhelm Hofmeister: The Origin of Causal-analytical Research in Plant Development." *American Journal of Botany* 83, no 12 (1996): 1647–1660. <https://doi.org/10.1002/j.1537-2197.1996.tb12823.x>

- Kay, Lily. *The Molecular Vision of Life: Caltech, the Rockefeller Foundation, and the Rise of the New Biology*. Oxford: Oxford University Press, 1993.
- Kay, Lily. "Who Wrote the Book of Life? Information and the Transformation of Molecular Biology, 1945–55." *Science in Context* 8, no. 4 (1995): 609–634. <https://doi.org/10.1017/S0269889700002210>
- Keller, Evelyn. *The Century of the Gene*. Cambridge: Harvard University Press, 2000.
- Kitcher, Philip. "1953 and All That: A Tale of Two Sciences." *Philosophical Review* 93, no 3 (1984): 335–373.
- Kitcher, Philip. "Battling the Undead: How (and How Not) to Resist Genetic Determinism." In *Thinking About Evolution: Historical, Philosophical and Political Perspectives*, edited by Rama Singh, Costas Krimbas, Diane Paul and John Beatty, 396–414. Cambridge: Cambridge University Press, 2001.
- Kitcher, Philip. *In Mendel's Mirror: Philosophical Reflections on Biology*. Oxford: Oxford University Press, 2003.
- Kjosavik, Frode. "From Symbolism to Information? Decoding the Gene Code." *Biology and Philosophy* 22, no 3 (2007): 333–349. <http://dx.doi.org/10.1007/s10539-006-9044-8>
- Kjosavik, Frode. "Genes, Structuring Powers and the Flow of Information in Living Systems." *Biology and Philosophy* 29, no 3 (2014): 379–394. <https://doi.org/10.1007/s10539-013-9407-x>
- Kohler, Robert. "The Management of Science: The Experience of Warren Weaver and the Rockefeller Foundation Programme in Molecular Biology." *Minerva* 14, no 3 (1976): 279–306. <https://doi.org/10.1007/bf01096274>
- Kumar, Lucy. "Information, Meaning, and Error in Biology." *Biological Theory* 9, no 1 (2014): 89–99.
- Kupferschmidt, Kai. "Biologist Raise Alarm over Changes to Biopiracy Rules." *Science* 361, no 6397 (2018): 14. <https://doi.org/10.1126/science.361.6397.14>
- Laird, Sarah, Rachel Wynberg, Michelle Rourke, Fran Humphries, Manuel Ruiz Muller and Charles Lawson. "Rethink the Expansion of Access and Benefit Sharing." *Science* 367, no 6483 (2020): 1200–1202. <https://doi.org/10.1126/science.aba9609>
- Lawson, Charles and Catherine Pickering. "Patent Laws Will Undermine Access Provisions in the Environment Protection and Biodiversity Conservation Bill 1998." *Environmental and Planning Law Journal* 15, no 6 (1998): 401–409.
- Lawson, Charles, Fran Humphries and Michelle Rourke. "The Future of Information under the CBD, Nagoya Protocol, Plant Treaty and PIP Framework." *Journal of World Intellectual Property* 22, no 3–4 (2019): 103–119. <https://doi.org/10.1111/jwip.12118>
- Lawson, Charles, Michelle Rourke and Fran Humphries. "Information as the Latest Site of Conflict in the Ongoing Contests about Access to and Sharing the Benefits from Exploiting Genetic Resources." *Queen Mary Journal of Intellectual Property* 10, no 1 (2020): 7–33. <https://doi.org/10.4337/qmjip.2020.01.01>
- Levy, Aaron. "Information in Biology: A Fictionalist Account." *Noûs* 45, no 4 (2011): 640–657. <https://doi.org/10.1111/j.1468-0068.2010.00792.x>
- Martins, L.A.-C.P. "Did Sutton and Boveri Propose the So-called Sutton-Boveri Chromosome Hypothesis?" *Genetics and Molecular Biology* 22, no 2 (1999): 261–272. <https://doi.org/10.1590/S1415-47571999000200022>
- Maynard-Smith, John. "The Concept of Information in Biology." *Philosophy of Science* 67, no 2 (2000): 177–194. https://doi.org/10.1007/978-94-017-0475-5_18
- Mekonnen, Dawit and David Spielman. "Changing Patterns in Genebank Acquisitions of Crop Genetic Materials: An Analysis of Global Policy Drivers and Potential Consequences." *Food Policy* 105 (2021): 102161. <https://doi.org/10.1016/j.foodpol.2021.102161>
- Mendel, Gregor. "Versuche über Pflanzenhybriden." *Verhandlungen des naturforschenden Vereines zu Brünn* 4 (1865): 3–47. For a translation and commentary by Staffan Müller-Wille and Kersten Hall (2016). British Society for the History of Science Translation Series. <http://www.bsbs.org.uk/bsbs-translations/mendel>
- Millikan, Ruth. *Language, Thought and Other Biological Categories*. Cambridge: MIT Press, 1984.
- Morange, Michel. *A History of Molecular Biology*. Cambridge: Harvard University Press, 1998.
- Morange, Michael. "What History Tells Us. XIII. Fifty Years of the Central Dogma." *Journal of Biosciences* 33, no 2 (2008): 171–175. <http://dx.doi.org/10.1007/s12038-008-0034-7>
- Morgan, T. "Sex-limited Inheritance in *Drosophila*." *Science* 32, no 812 (1910): 120–122. <https://doi.org/10.1126/science.32.812.120>
- Morgan, T., A. Sturtevant, H. Muller and C. Bridges. *The Mechanism of Mendelian Heredity*. New York: Henry Holt, 1915.
- Morris, K. and J. Mattick. "The Rise of Regulatory RNA." *Nature Reviews Genetics* 15, no 6 (2014): 423–437. <https://dx.doi.org/10.1038%2Fnr3722>
- Moss, Lenny. *What Genes Can't Do*. Cambridge: MIT Press, 2003.
- Muller, Michael, Lennart Martens and Rolf Apweiler. "Annotating the Human Proteome: Beyond Establishing a Parts List." *Biochimica et Biophysica Acta* 1774, no 2 (2007): 175–191.
- Neumann, D., A. Borisenko, J. Coddington, C. Häuser, C. Butler, A. Casino et al. "Global Biodiversity Research Tied Up by Juridical Interpretations of Access and Benefit Sharing." *Organisms Diversity and Evolution* 18, no 1 (2018): 1–12. <http://dx.doi.org/10.1007%2Fs13127-017-0347-1>

- Neumann-Held, Eva. “Genes – Causes – Codes: Deciphering DNA’s Ontological Privilege.” In *Genes in Development: Re-reading the Molecular Paradigm*, edited by Eva Neumann-Held and Christoph Rehmann-Sutter, 238–271. Durham, NC: Duke University Press, 2006.
- Nijar, Gurdial Singh, Sélim Louafi and Eric Welch. “The Implementation of the Nagoya ABS Protocol for the Research Sector: Experience and Challenges.” *International Environmental Agreements: Politics, Law and Economics* 17, no 5 (2017): 607–621.
- Olby, Robert. “Mendel no Mendelian?” *History of Science* 17, no 1 (1979): 53–72.
<https://doi.org/10.1177%2F007327537901700103>
- Oyama, Susan. “Causal Democracy and Causal Contributions in Developmental Systems Theory.” *Philosophy of Science* 67 (2000): S332–S347. <https://doi.org/10.1086/392830>
- Painter, Theophilus. “A New Method for the Study of Chromosome Aberrations and the Blotting of Chromosome Maps in *Drosophila melanogaster*.” *Genetics* 19, no 3 (1934): 175–188. <https://doi.org/10.1093/genetics/19.3.175>
- Pocheville, Arnaud, Paul Griffiths and Karola Stotz. “Comparing Causes – An Information-theoretic Approach to Specificity, Proportionality and Stability.” In *Logic, Methodology and Philosophy of Science: Proceedings of the 15th International Congress*, edited by Hannes Leitgeb, Ilkka Niiniluoto and Paivi Seppala, 93–102. London: College Publications, 2017.
- Pocheville, Arnaud. “Biological Information as Choice and Construction.” *Philosophy of Science* 85, no 5 (2018): 1012–1025. <https://doi.org/10.1086/699832>
- Portin, Petter and Adam Wilkins. “The Evolving Definition of the Term ‘Gene.’” *Genetics* 205, no 4 (2017): 1353–1364. <https://doi.org/10.1534/genetics.116.196956>
- Prathapan, Divakaran, Rohan Pethiyagoda, Kamaljit Bawa, Peter Raven and Priyadarsanan Dharma Rajan. “When the Cure Kills – CBD Limits Biodiversity Research.” *Science* 360, no 6396 (2018): 1405–1406. <https://doi.org/10.1126/science.aat9844>
- Quastler, Henry. *Essays on the Use of Information Theory in Biology*. Urbana: University of Illinois Press, 1953.
- Radick, Gregory. “Making Sense of Mendelian Genes.” *Interdisciplinary Science Reviews* 45, no 3 (2020): 299–314. <https://doi.org/10.1080/03080188.2020.1794387>
- Ranganathan, Shoba, Kenta Nakai and Christian Schonbach, eds. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*. Amsterdam: Elsevier, 2018.
- Robert, Jason Scott. *Embryology, Epigenesis and Evolution*. Cambridge: Cambridge University Press, 2004.
- Roll-Hansen, Nils. “Sources of Wilhelm Johannsen’s Genotype Theory.” *Journal of the History of Biology* 42, no 3 (2009): 457–493. <https://doi.org/10.1007/s10739-008-9166-8>
- Román-Roldán, Ramón, Pedro Bernaola-Galván and José Oliver. “Application of Information Theory to DNA Sequence Analysis: A Review.” *Pattern Recognition* 9, no 7 (1996): 1187–1194. [https://doi.org/10.1016/0031-3203\(95\)00145-X](https://doi.org/10.1016/0031-3203(95)00145-X)
- Rosenberg, Alex. *Darwinian Reductionism: Or, How to Stop Worrying and Love Molecular Biology*. Chicago: University of Chicago Press, 2006.
- Rosenberg, Alex. “Is Epigenetic Inheritance a Counterexample of the Central Dogma?” *History and Philosophy of the Life Sciences* 28, no 4 (2006): 549–565.
- Rourke, Michelle, Mark Eccleston-Turner, Alexandra Phelan and Lawrence Gostin. “Policy Opportunities to Enhance Sharing for Pandemic Research.” *Science* 368, no 6492 (2020): 716–718. <https://doi.org/10.1126/science.abb9342>
- Rout, Saroj, Michael Friedmann, Roland Riek and Jason Greenwald. “A Prebiotic Template-directed Peptide Synthesis Based on Amyloids.” *Nature Communications* 9 (2018): 234. <https://doi.org/10.1038/s41467-017-02742-3>
- Roux, Wilhelm. *Ueber die Bedeutung der Kerntheilungsfiguren: Eine Hypothesische Erörterung*. Leipzig: Verlag von Wilhelm Engelmann, 1883.
- Sanger, F., A. Coulson, T. Friedmann, G. Air, B. Barrell, N. Brown et al. “The Nucleotide Sequence of Bacteriophage ϕ X174.” *Journal of Molecular Biology* 125, no 2 (1978): 225–246. [https://doi.org/10.1016/0022-2836\(78\)90346-7](https://doi.org/10.1016/0022-2836(78)90346-7)
- Sara, Rodrigo, Markus Wyss, René Custers, Anouk in ’t Veld and Dominic Muyldermans. “A Need for Recalibrating Access and Benefit Sharing.” *EMBO Reports*, e53973 (2021): 1–4. <https://doi.org/10.15252/embr.202153973>
- Sarabhai, A., A. Stretton, S. Brenner and A. Bolle. “Co-linearity of the Gene with the Polypeptide Chain.” *Nature* 201, no 4914 (1964): 13–17.
- Sarkar, Sahotra. “Decoding ‘Coding’-Information and DNA.” *BioScience* 46, no 11 (1996): 857–864. <https://doi.org/10.2307/1312971>
- Sarkar, Sahotra. “Biological Information: A Skeptical Look at Some Central Dogmas of Molecular Biology.” In *The Philosophy and History of Molecular Biology: New Perspectives*, edited by Sahotra Sarkar, 187–232. Dordrecht: Kluwer, 1996.
- Sarkar, Sahotra. “Genes Encode Information for Phenotypic Traits.” In *Contemporary Debates in Philosophy of Science*, edited by Christopher Hitchcock, 259–274. Malden: Blackwell Publishing, 2003.
- Schaffner, Kenneth. “The Watson-Crick Model and Reductionism.” *British Journal for the Philosophy of Science* 20, no 4 (1969): 325–348. <https://doi.org/10.1093/bjps/20.4.325>

- Schmucker, Dietmar and Brian Chen. "Dscam and DSCAM: Complex Genes in Simple Animals, Complex Animals Yet Simple Genes." *Genes and Development* 23, no 2 (2009): 147–156.
- Schrödinger, Erwin. *What is Life? The Physical Basis of the Living Cell*. Cambridge: Cambridge University Press, 1944.
- Scott, Deborah and Dominic Berry. "Genetic Resources in the Age of the Nagoya Protocol and Gene/Genome Synthesis." Report and Analysis of an Interdisciplinary Workshop: University of Edinburgh, 2017.
https://haseloff.plantsci.cam.ac.uk/resources/SynBio_reports/Nagoya_workshop_report_2018.pdf
- Shannon, Claude. "A Mathematical Theory of Communication." *Bell System Technical Journal* 27, no 3 (1948): 379–423.
<https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Shannon, Claude and Warren Weaver. *The Mathematical Theory of Communication*. Urbana: University of Illinois Press, 1949.
- Shapiro, Jim, Lorne Machattie, Larry Eron, Garret Ihler, Karin Ippen and Jon Beckwith. "Isolation of Pure *lac* Operon DNA." *Nature* 224, no 5221 (1969): 768–774.
- Shea, Nicholas. "Representation in the Genome and in Other Inheritance Systems." *Biology and Philosophy* 22, no 3 (2007): 313–331. <https://doi.org/10.1007/s10539-006-9046-6>
- Shea, Nicholas. "Consumers Need Information: Supplementing Teleosemantics with an Input Condition." *Philosophy and Phenomenological Research* 75, no 2 (2007): 404–435. <https://doi.org/10.1111/j.1933-1592.2007.00082.x>
- Shea, Nicholas. "Developmental Systems Theory Formulated as a Claim about Inherited Representations." *Philosophy of Science* 78, no 1 (2011): 60–82. <https://doi.org/10.1086/658110>
- Shea, Nicholas. "Inherited Representations are Read in Development." *British Journal for the Philosophy of Science* 64, no 1 (2013): 1–31. <https://doi.org/10.1093/bjps/axr050>
- Smith, David, Harriet Hinz, Joseph Mulema, Philip Weyl and Matthew Ryan. "Biological Control and the Nagoya Protocol on Access and Benefit Sharing – A Case of Effective Due Diligence." *Biocontrol Science and Technology* 28, no 10 (2018): 914–926. <https://doi.org/10.1080/09583157.2018.1460317>
- Stegmann, Ulrich. "The Arbitrariness of the Genetic Code." *Biology and Philosophy* 19, no 2 (2004): 205–222.
<https://doi.org/10.1023/B:BIPH.0000024412.82219.a6>
- Stegmann, Ulrich. "Genetic Information as Instructional Content." *Philosophy of Science* 72, no 3 (2005): 425–443.
<https://doi.org/10.1086/498472>
- Stegmann, Ulrich. "DNA, Inference and Information." *British Journal for the Philosophy of Science* 60, no 1 (2009): 1–17.
<https://doi.org/10.1093/bjps/axn041>
- Stegmann, Ulrich. "Varieties of Parity." *Biology and Philosophy* 27, no 6 (2012): 903–918.
- Sterelny, Kim. "The 'Genetic Program' Program: A Commentary on Maynard-Smith on Information in Biology." *Philosophy of Science* 67, no 2 (2000): 195–201.
- Sterelny, Kim and Paul Griffiths. *Sex and Death: An Introduction to the Philosophy of Biology*. Chicago, Illinois: University of Chicago Press 1999.
- Sterelny, Kim and Philip Kitcher. "The Return of the Gene." *Journal of Philosophy* 85, no 7 (1988): 339–361.
<https://doi.org/10.2307/2026953>
- Sterelny, Kim, Kelly Smith and Michael Dickison. "The Extended Replicator." *Biology and Philosophy* 11, no 3 (1996): 377–403.
- Stotz, Karola. "With 'Genes' Like That, Who Needs an Environment? Postgenomics's Argument for the 'Ontogeny of Information.'" *Philosophy of Science* 73, no 5 (2006): 905–917. <https://doi.org/10.1086/518748>
- Stotz, Karola. "Molecular Epigenesis: Distributed Specificity as a Break in the Central Dogma." *History and Philosophy of the Life Sciences* 28, no 4 (2006): 533–548.
- Stotz, Karola and Paul Griffiths. "Biological Information, Causality and Specificity – An Intimate Relationship." In *From Matter to Life: Information and Causality*, edited by Sara Imari Walker, Paul Davies and George Ellis, 366–390. Cambridge: Cambridge University Press, 2017.
- Sturtevant, A. "The Linear Arrangement of the Six Sex Linked Factors in *Drosophila*, as Shown by their Mode of Association." *Journal of Experimental Zoology* 14, no 1 (1913): 43–59. <https://doi.org/10.1002/jez.1400140104>
- Šustar, Predrag. "Crick's Notion of Genetic Information and the 'Central Dogma' of Molecular Biology." *British Journal for the Philosophy of Science* 58, no 1 (2007): 13–24. <https://doi.org/10.1093/bjps/axl018>
- Sutton, Walter. "The Chromosomes in Heredity." *Biological Bulletin* 4, no 5 (1903): 231–250.
<https://doi.org/10.2307/1535741>
- Tang, Wei, Yongjun Fei and Michael Page. "Biological Significance of RNA Editing in Cells." *Molecular Biotechnology* 59, no (2012): 91–100. <https://doi.org/10.1007/s12033-012-9498-7>
- Temin, Howard and Satoshi Mizutani. "RNA-dependent DNA Polymerase in Virions of Rous sarcoma Virus." *Nature* 226, no 5252 (1970): 1211–1213.
- Thieffry, Denis and Sahotra Sarkar. "Forty Years under the Central Dogma." *Trends in Biochemical Science* 23, no 8 (1998): 312–316. [https://doi.org/10.1016/s0968-0004\(98\)01244-4](https://doi.org/10.1016/s0968-0004(98)01244-4)

- Venter, Craig, Mark Adams, Eugene Myers, Peter Li, Richard Mural and Granger Sutton. "The Sequence of the Human Genome." *Science* 291, no 5507 (2001): 1304–1351. <https://doi.org/10.1126/science.1058040>
- Wagner, Andreas. "The Role of Robustness in Phenotypic Adaptation and Innovation." *Proceedings of the Royal Society B: Biological Sciences* 279, no 1732 (2012): 1249–1258. <https://doi.org/10.1098/rspb.2011.2293>
- Wanscher, Johan. "An Analysis of Wilhelm Johannsen's Genetical Term 'Genotype' 1909–26." *Hereditas* 79, no 1 (1975): 1–4. <https://doi.org/10.1111/j.1601-5223.1975.tb01456.x>
- Ward Bynum, Terrell. "Informational Metaphysics: The Informational Nature of Reality." In *The Routledge Handbook of Philosophy of Information*, edited by Luciano Floridi, 219–234. London: Taylor & Francis Group, 2016.
- Watanabe, Myrna. "The Nagoya Protocol: The Conundrum of Defining Digital Sequence Information." *BioScience* 69, no 6 (2019): 480. <https://doi.org/10.1093/biosci/biz034>
- Waters, Kenneth. "Genes Made Molecular." *Philosophy of Science* 61, no 2 (1994): 163–185. <https://doi.org/10.1086/289794>
- Waters, Kenneth. "Causes that Make a Difference." *Journal of Philosophy* 104, no 11 (2007): 551–579.
- Watson, James and Friedrich Crick. "Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid." *Nature* 171, no 4356 (1953): 737–738.
- Watson, James and Friedrich Crick. "Genetical Implications of the Structure of Deoxyribose Nucleic Acid." *Nature* 171, no 4361 (1953): 964–967.
- Weaver, Warren. "Molecular Biology: Origin of the Term." *Science* 170, no 3958 (1970): 581–582. <https://doi.org/10.1126/science.170.3958.581-a>
- Weber, Marcel and Michael Ruse. *Philosophy of Experimental Biology*. Cambridge: Cambridge University Press, 2004.
- Weber, Marcel. "The Central Dogma as a Thesis of Causal Specificity." *History and Philosophy of the Life Sciences* 28, no 4 (2006): 595–609.
- Weber, Marcel. "Which Kind of Causal Specificity Matters Biologically?" *Philosophy of Science* 84, no 3 (2017): 574–585.
- Wilhelm Johannsen. *Elemente der exakten Ererblichkeitslehre (Elements of an Exact Theory of Heredity)*. Jena: Gustav Fischer, 1909.
- Williams, China, Alexandra Walsh, Valentina Vaglica, Aysegül Sirakaya, Manuela da Silva and Gemedo Dalle. "Conservation Policy: Helping or Hindering Science to Unlock Properties of Plants and Fungi." *Plants, People, Planet* 2, no 5 (2020): 535–545. <https://doi.org/10.1002/ppp3.10139>
- Williams, George. *Plan and Purpose in Nature: Science Masters*. London: Weidenfeld & Nicolson, 1992.
- Witty, Michael. "Pollen Development, Membranes and Features of the Nucleus in *Tradescantia* and related genera: A Translation of Wilhelm Hofmeister's 1848 Paper 'Ueber die Entwicklung des Pollens.'" *Huntia* 15, no 2 (2015): 75–86.
- Witzany, Günther. *Life: The Communicative Structure – A New Philosophy of Biology*. Witzany: Norderstedt, 2000.
- Woodward, James. *Making Things Happen*. Oxford: Oxford University Press, 2003.
- Wright, Kate, Nick Fisk and Dina Newman. "DNA → RNA: What Do Students Think the Arrow Means?" *Life Sciences Education* 13, no. 2 (2014): 338–348. <https://doi.org/10.1187/cbe.cbe-13-09-0188>
- Wyatt, H. "When Does Information Become Knowledge." *Nature* 235, no 5333 (1972): 86–89.
- Yčas, Martynas. "The Protein Text." In *Symposium on Information Theory in Biology*, edited by Hubert Yockey, 70–100. New York: Pergamon Press, 1956.
- Yockey, H. "An Application of Information Theory to the Central Dogma and the Sequence Hypothesis." *Journal of Theoretical Biology* 46, no 2 (1974): 369–406. [https://doi.org/10.1016/0022-5193\(74\)90005-8](https://doi.org/10.1016/0022-5193(74)90005-8)
- Young, Tomme and Morten Tvedt. *Drafting Successful Access and Benefit-sharing Contracts*. Leiden Netherlands: Brill, 2017.
- Yoxen, E. "Where Does Schroedinger's 'What is Life?' Belong in the History of Molecular Biology?" *History of Science* 17, no 1 (1979): 17–52. <https://doi.org/10.1177%2F007327537901700102>

Primary Materials

India

Biological Diversity Act 2002 (India).

Kenya

Environmental Management and Co-ordination (Conservation on Biological Diversity and Resources, Access to Genetic Resources and Benefit Sharing) Regulations 2006 (Kenya).

Malaysia

Access to Biological Resources and Benefit-Sharing Act 2017 (Malaysia).

International

Convention on Biological Diversity (opened for signature 5 June 1992, 1760 UNTS 79 [entered into force 29 December 1993]).
 Marrakesh Agreement Establishing the World Trade Organization, Annex 1C (Agreement on Trade Related Aspects of Intellectual Property Rights) (signed 15 April 1994, 1867 UNTS 154 [entered into force 1 January 1995]).
 Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity, (adopted 29 October 2010, 3008 UNTS 1 [entered into force 12 October 2014]).

Secondary Legal Materials

Contracting Party Responses to the DSI Debates

African Union Commission on behalf of the African Group. <https://www.cbd.int/abs/DSI-views/2019/AfricanGroup-DSI.pdf>
 Australia. <https://www.cbd.int/abs/DSI-views/2019/Australia-DSI.pdf>
 Brazil. <https://www.cbd.int/abs/DSI-views/2019/Brazil-DSI.pdf>
 Canada. <https://www.cbd.int/abs/DSI-views/2019/Canada-DSI.pdf>
 Ethiopia. <https://www.cbd.int/abs/DSI-views/2019/Ethiopia-DSI.pdf>
 European Union and its Member States. <https://www.cbd.int/abs/DSI-views/2019/EU-MS-DSI.pdf>
 India. <https://www.cbd.int/abs/DSI-views/2019/India-DSI.pdf>
 Iran. <https://www.cbd.int/abs/DSI-views/2019/Iran-DSI.pdf>
 Japan. <https://www.cbd.int/abs/DSI-views/2019/Japan-DSI.pdf>
 Korea. <https://www.cbd.int/abs/DSI-views/2019/ROK-DSI.pdf>
 Madagascar. <https://www.cbd.int/abs/DSI-views/2019/Madagascar-DSI.pdf>
 South Africa. <https://www.cbd.int/abs/DSI-views/2019/SouthAfrica-DSI.pdf>
 Switzerland. <https://www.cbd.int/abs/DSI-views/2019/Switzerland-DSI.pdf>

Governing Body Documents

CBD/COP/13/25 (2016) Conference of the Parties to the Convention on Biological Diversity, *Report of the Conference of the Parties to the Convention on Biological Diversity on its Thirteenth Meeting*. Montreal: Secretariat of the Convention on Biological Diversity.
 CBD/COP/14/14 (2018) Conference of the Parties to the Convention on Biological Diversity, *Report of the Conference of the Parties to the Convention on Biological Diversity on its Fourteenth Meeting*. Montreal: Secretariat of the Convention on Biological Diversity.
 CBD/DSI/AHTEG/2018/1/3 (2018) *Ad Hoc* Technical Expert Group on Digital Sequence Information on Genetic Resources, *Fact-Finding and Scoping Study on Digital Sequence Information on Genetic Resources in the Context of the Convention on Biological Diversity and the Nagoya Protocol*. Montreal: Secretariat of the Convention on Biological Diversity.
 CBD/DSI/AHTEG/2018/1/4 (2018) *Ad Hoc* Technical Expert Group on Digital Sequence Information on Genetic Resources, *Report of the Ad Hoc Technical Expert Group on Digital Sequence Information on Genetic Resources*. Montreal: Secretariat of the Convention on Biological Diversity.
 CBD/DSI/AHTEG/2020/1/3 (2018) *Ad Hoc* Technical Expert Group on Digital Sequence Information on Genetic Resources, *Digital Sequence Information on Genetic Resources: Concept, Scope and Current Use*. Montreal: Secretariat of the Convention on Biological Diversity.
 CBD/DSI/AHTEG/2020/1/4 (2020) *Ad Hoc* Technical Expert Group on Digital Sequence Information on Genetic Resources, *Combined Study on Digital Sequence Information (DSI) in Public and Private Databases and Traceability*. Montreal: Secretariat of the Convention on Biological Diversity.
 CBD/DSI/AHTEG/2020/1/5 (2020) *Ad Hoc* Technical Expert Group on Digital Sequence Information on Genetic Resources, *Fact-finding Study on How Domestic Measures Address Benefit-sharing Arising from Commercial and Non-commercial Use of Digital Sequence Information on Genetic Resources and Address the Use of Digital Sequence Information on Genetic Resources for Research and Development*. Montreal: Secretariat of the Convention on Biological Diversity.
 CBD/DSI/AHTEG/2020/1/7 (2020) *Ad Hoc* Technical Expert Group on Digital Sequence Information on Genetic Resources, *Report of the Ad Hoc Technical Expert Group on Digital Sequence Information on Genetic Resources*. Montreal: Secretariat of the Convention on Biological Diversity.
 CBD/NP/MOP/2/13 (2016) Conference of the Parties to the Convention on Biological Diversity Serving as the Meeting of the Parties to the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of the Benefits Arising from their Utilization, *Report of the Conference of the Parties to the Convention on Biological Diversity Serving as the Meeting of the Parties to the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of the Benefits Arising from their Utilization on its Second Meeting*. Montreal: Secretariat of the Convention on Biological Diversity.

- CBD/NP/MOP/3/10 (2018) Conference of the Parties to the Convention on Biological Diversity Serving as the Meeting of the Parties to the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of the Benefits Arising from their Utilization, *Report of the Conference of the Parties to the Convention on Biological Diversity Serving as the Meeting of the Parties to the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of the Benefits Arising from their Utilization on its Third Meeting*. Montreal: Secretariat of the Convention on Biological Diversity.
- CBD/SBSTTA/22/12 (2018) Subsidiary Body on Scientific, Technical and Technological Advice, *Report of the Subsidiary Body on Scientific, Technical and Technological advice on its Twenty-Second Meeting*. Montreal: Secretariat of the Convention on Biological Diversity.
- CBD/WG2020/3/4 (2021) Open-Ended Working Group on the Post-2020 Global Biodiversity Framework, *Digital Sequence Information on Genetic Resources*. Montreal: Secretariat of the Convention on Biological Diversity.
- UNEP/CBD/COP/3/20 (1996) Conference of the Parties to the Convention on Biological Diversity, *Access to Genetic Resources*. Montreal: Secretariat of the Convention on Biological Diversity.
- UNEP/CBD/COP/12/6 (2014) Open-Ended Ad Hoc Intergovernmental Committee for the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization, *Report of the Third Meeting of the Open-Ended Ad Hoc Intergovernmental Committee for the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization*. Montreal: Secretariat of the Convention on Biological Diversity.
- UNEP/CBD/COP/12/11 (2014) Conference of the Parties to the Convention on Biological Diversity, *Progress Report on Technical and Scientific Cooperation and the Clearing House Mechanism*. Montreal: Secretariat of the Convention on Biological Diversity.
- UNEP/CBD/ICNP/3/6 (2014) Open-Ended Ad Hoc Intergovernmental Committee for the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization, *Report on Progress in the Implementation of the Pilot Phase of the Access and Benefit Sharing Clearing House*. Montreal: Secretariat of the Convention on Biological Diversity.
- UNEP/CBD/WG-ABS/7/2 (2008) Ad Hoc Open-Ended Working Group on Access and Benefit Sharing, *Report of the Meeting of the Group of Legal and Technical Experts on Concepts, Terms, Working Definitions and Sectoral Approaches*. Montreal: Secretariat of the Convention on Biological Diversity.
- UNEP/CBD/WG-ABS/9/INF/1 (2010) Ad Hoc Open-Ended Working Group on Access and Benefit Sharing, *The Concept of 'Genetic Resources' in the Convention on Biological Diversity and How It Relates to a Functional International Regime on Access and Benefit Sharing*. Montreal: Secretariat of the Convention on Biological Diversity.
- UNEP/CBD/WGRI/5/3/Add.2 (2014) Ad Hoc Open-Ended Working Group on Review of Implementation of the Convention, *Progress Report on the Clearing House Mechanism*. Montreal: Secretariat of the Convention on Biological Diversity.
- WHA64.5 (2011) World Health Organisation, *Pandemic Influenza Preparedness: Sharing of Influenza Viruses and Access to Vaccines and Other Benefits*, Sixty-Fourth World Health Assembly. Geneva: World Health Organisation.

Australia

Department of Environment and Energy, *Biodiversity*. <http://www.environment.gov.au/biodiversity>