The Central Dogma Is Empirically Inadequate ... No Matter How We Slice It

M. Polo Camacho*

Roughly, the Central Dogma of molecular biology states that DNA codes for protein, not the other way around. This principle, which is still heralded in biology as an important element of contemporary biological theory (Raineri 2001; Morris 2013), has received much critical attention since its original formulation by Francis Crick in 1958 (Crick 1958; Crick 1970; Maynard Smith 1993; Morris 2013; Rosenberg 2006; Weber 2006; Dawkins 1982; Sarkar 1996; Sarkar 2004; Watson 1965; Raineri 2001; Šustar 2007). Yet, despite the vast criticism, there is much about the Dogma that has not been said. Existing discussions, for example, gloss over the many distinct, logically independent readings of the Central Dogma that have been defended in the philosophical and biological literature. This oversight makes it difficult to see which dogma is being criticized, and, more generally, what the overall upshot of these discussions should be taken to be. My aim is to fix this. To this end, I first examine five logically distinct interpretations of the Central Dogma—some of which have been overlooked in the literature to date—and then make explicit why each of these interpretations fail. I conclude that the Central Dogma is empirically inadequate no matter how we slice it.

Keywords

The Central Dogma • biological information • causal specificity • Francis Crick • genetic code • inheritance of acquired traits

1 Introduction

Roughly, the Central Dogma of molecular biology states that DNA codes for protein, not the other way around. This principle, which is still heralded as an important element of contemporary biological theory (Raineri 2001; Morris 2013), has received much critical attention since its original formulation by Francis Crick in 1958 (Crick 1958; Crick 1970; Maynard Smith 1993; Morris 2013; Rosenberg 2006; Weber 2006; Dawkins 1982; Sarkar 1996, 2004; Watson 1965;

*Department of Philosophy, University of Kansas, 1445 Jayhawk Blvd., Wescoe Hall, Room 3108, Lawrence, Kansas, 66045 USA, mpcamacho@ku.edu

Received 10 October 2018; Revised 04 February 2019; Accepted 31 March 2019
doi:10.3998/ptpbio.16039257.0011.006

Raineri 2001; Šustar 2007). Some have argued that the principle should be rejected, on the grounds that it fails to fully capture the ins-and-outs of protein synthesis (Stotz and Griffiths 2013; Stotz 2006), while others have argued that the Dogma is predicated on notions of information that are simply implausible (Sarkar 1996). Yet, despite all this criticism, there is much about the Dogma that *has not* been said. Existing discussions, for example, gloss over the many distinct, logically independent readings of the Central Dogma that have been defended in the philosophical and biological literature, making it difficult to see which dogma is being criticized. Additionally, this oversight makes it unclear what the overall upshot of these discussions should be taken to be. My aim in this paper is to fix this.

In what follows, I set out a comprehensive overview of five different conceptions of the Central Dogma and assess them for their plausibility. Unlike previous discussions, this paper explores logically distinct interpretations of the Central Dogma—some of which have been overlooked in the literature to date—and makes explicit why each of these interpretations fail. The upshot then is the most up-to-date and overarching empirical assessment of the Central Dogma in the literature. I conclude that the Central Dogma is problematic no matter how we slice it. I want to emphasize from the get-go that my goal here is only to assess the *empirical adequacy* of the Central Dogma. Of course, it may be the case that the dogma is useful and defensible for other reasons—after all, even empirically inadequate claims can still be predictive and useful. However, whether this is so needs to be assessed on a different occasion. (I revisit this point below.)

I begin this paper by engaging with, and offering a brief overview of, interpretations of the Dogma that have appeared in the philosophical literature. Partly drawing on contemporary arguments, and partly by advancing novel arguments, it is then shown that these interpretations fail to be plausible. In the sections that follow, I explore, and ultimately refute, the largely overlooked interpretations of Francis Crick (Crick 1958) and John Maynard Smith (Smith 1993).

2 Five Dogmas of Molecular Biology

Since its formulation, the Central Dogma has been viewed as a fundamental principle of biology, providing a firm foundation for the life sciences. Bruno Strasser notes, for example, how the Central Dogma represented a turning point in molecular science, as it came to replace three-dimensional models of protein synthesis, and in so doing “brought the problem of gene action and protein synthesis down to one dimension” (Strasser 2006, 506). Michel Morange notes how the Dogma’s central claims concerning the relationship between DNA, RNA and protein were shaped by evolutionary history (Morange 2008, 2009). Many others have gone as far as to claim that the Dogma inherently denies the inheritance of acquired traits, and with it Lamarckian conceptions of evolution (Maynard Smith 1993; Judson 1979; Dawkins 1982; Cobb 2017; Wilkins 2002). In a 1958 article titled “On Protein Synthesis,” Francis Crick, the Dogma’s founder, once claimed that it would be “an instructive exercise” to try and build a theory about protein synthesis without the Dogma; such an attempt, Crick argued, would inevitably leave one “in the wilderness.”

An immediate problem for assessments of the plausibility of these claims—and the Dogma more generally—is the fact that the *Central Dogma* is not a unitary thesis with widely accepted meaning. The Dogma’s interpretations abound, making it difficult, without first specifying exactly which interpretation we are talking about, to demonstrate how exactly the Central Dogma is problematic. Hence, it is necessary to make explicit which version of the Dogma we are talking about. In the paragraphs that follow, I will sketch five formulations of the Dogma very

briefly, since the point is simply to provide enough information to make the analysis easier to follow. Each interpretation will be spelled out in more detail in the sections below.

The first version of the Central Dogma—hereafter, CD₁—says that DNA specifies RNA, which specifies protein. This view bears a close resemblance to a formulation often attributed to James Watson, which states roughly that DNA makes RNA makes protein (Watson 1965). Paul Griffiths and Karola Stotz sum up this view nicely: “the linear sequence of nucleotides in a segment of a DNA molecule specifies the linear sequence of nucleotides in an RNA molecule, and that molecule in turn determines the linear sequence of amino acids in a protein” (Griffiths and Stotz 2013, 49). CD₁ has also figured prominently in introductory textbooks in molecular biology. Morris (2013) writes:

How is the sequence of amino acids specified? It is specified by the sequence of nucleotides in the DNA, in coded form. The decoding of the information takes place according to the Central Dogma of molecular biology, which defines information flow in a cell from DNA to RNA to protein In transcription, the sequence of bases along part of a DNA strand is used as a template in the synthesis of a complementary sequence of bases in a molecule of RNA In translation, the sequence of bases in an RNA molecule known as messenger RNA (mRNA) is used to specify the order in which successive amino acids are added to a newly synthesized polypeptide chain. (77)

To better illustrate CD₁, consider the DNA sequence AGG, which is made up of nucleotides Adenine, Guanine, and Guanine. According to CD₁, this sequence is transcribed into nucleotides in RNA segments, which are then translated into the codon Arginine. It is in this linear sense that DNA specifies protein.

The second version of the Dogma—hereafter, CD₂—conforms with a kind of “geno-centrism” (Rosenberg 2006; Weber 2006), in that it regards DNA as the most significant cause contributing to protein synthesis. Pedrag Sustar sums up this interpretation as follows: “Crick’s notion of genetic information can be defined in terms of a certain type of causal relation between a sequence of bases in the nucleic acids and a sequence of amino acids in the protein. We have seen that the DNA molecule has been assigned a privileged causal role in the relation in question” (Sustar 2007, 10). Note that this interpretation does not say that DNA is the *only* cause involved in protein synthesis, and therefore does not exclude other causes, such as transcription proteins or enzymes, as significant factors involved in protein synthesis. CD₂ *just* says that DNA is the *most significant*—in a sense to be made precise in Section 3.2—cause of protein synthesis. This is different from the CD₁, in that CD₂ centers on the notion of causal efficacy and significance, rather than linear sequence specificity.

The third version of the Central Dogma—hereafter, CD₃—concerns the transfer of information. This version says that DNA alone carries information for protein. Since there are numerous informational interpretations of the Central Dogma in the philosophical literature, I restrict my focus to Sahotra Sarkar’s account which appeals to the notion of “semiotic information” (Sarkar 1996). This notion, for Sarkar, sets out the conditions under which DNA alone carries information for protein (Sarkar 1996, 2004).

The fourth version of the Central Dogma—hereafter CD₄—differs from the previous interpretations in that it is not a thesis about the linear specificity of protein via DNA, or about the causal significance of DNA, or about information flow from DNA to protein. CD₄, rather, bears a very close resemblance to Francis Crick’s original formulation of the Central Dogma. Crick distinguished between two principles: The Sequence Hypothesis and the Central Dogma, and noted that the Central Dogma “is not the same, as is commonly assumed, as the sequence

hypothesis In particular the sequence hypothesis was a positive statement, saying that the (overall) transfer nucleic acid-protein did exist, whereas the Central Dogma was a negative statement, saying that transfers from protein did not exist" (Crick 1970, 562). Put differently, for Crick, the Central Dogma says that the transfer of information from protein to protein and from protein to DNA is not possible; "once 'information' has passed into protein *it cannot get out again*" (Crick 1958). In short, while Watson's dogma—CD₁—tells us what DNA *can* do, Crick's dogma stresses what proteins *cannot* do. This version of the Central Dogma has seen little discussion in the literature to date (see also Graur 2019 on related issues).

The final version of the Central Dogma—hereafter CD₅—concerns the heritability of traits, and has been interpreted by numerous commentators as synonymous with, or at the very least implying, a rejection of the inheritance of acquired traits (Maynard Smith 1993; Dawkins 1983; Wilkins 2002). This conception of the Central Dogma says that genes cannot be modified by environmental factors in a way that produces heritable traits in the organism. As Dawkins (1982) puts it,

To regard an organism as a replicator ... is tantamount to a violation of the 'central dogma' of the non-inheritance of acquired characteristics It is wrong to say that 'just as genes can pass on their structure in gene lineages, organisms can pass on their structure in organism lineages.' (97)

Like the CD₄, the CD₅ has not been fully assessed in the context of the Central Dogma.

Now that all five dogmas are on the table, a few important notes are in order. The first is a note about logical independence. Specifically, CD₁ through CD₅ are logically independent theses. One might, for example, claim that DNA specifies protein, and deny that genes are the most significant causes involved in protein synthesis. Or, one might deny that DNA alone specifies protein, and yet claim that DNA *alone* carries information to protein; in other words, one might argue that a host of transcriptional and translational machinery specificity of protein, while consistently maintaining that information flows from DNA to protein. That CD₁ through CD₅ are logically independent implies not only that the Central Dogma has many formulations (further supporting the idea that the Dogma is not a unitary thesis), but also that the criticism of one Dogma does not necessarily imply the criticism of another. CD₂ is not necessarily made false, for example, by the falsity of CD₁. This note will be important to keep in mind as we evaluate each interpretation of the Central Dogma in the sections below.

The second point to note concerns the generality of the Dogma. Because the Central Dogma is meant to capture a general biological process, CD₁ through CD₅ should be empirically adequate, meaning that each formulation should be true in all cases of protein synthesis. The following quote from Crick conveys this point nicely:

These are the three transfers which the Central Dogma postulates *never* occur:

Protein → Protein

Protein → DNA

Protein → RNA

... the discovery of just *one* type of present day cell which could carry out any of the unknown transfers would shake the whole intellectual basis of molecular biology, and it is for this reason that the Central Dogma is as important today as when it was first proposed. (Crick 1970, emphasis added)

Here, Crick discusses bio-molecular interactions which "never occur": interactions from protein to protein, protein to RNA, and protein to DNA. The *only* interactions which do occur are those from DNA to DNA, DNA to RNA, and RNA to protein (Crick 1970). Given this, we

might say that if there is a particular empirical instance of bio-molecular interaction, which the Central Dogma stipulates never occurs, then we have grounds for rejecting it. More generally, if a particular interpretation of the Central Dogma amounts to a set of claims about fundamental bio-molecular interactions, and there are empirical instances in which demonstrate that such interactions do not occur, then we have grounds for rejecting the said interpretation. The motivation behind this “generality condition” is that we want a principle that captures significant aspects of protein synthesis (and biomolecular interactions) *generally*—i.e., we want to capture features that are true of all instances of protein synthesis. Given this, finding counterexamples—even if few in number—call the plausibility of the (relevant version of the) Central Dogma into question. This note will become important again below.

The third point concerns arbitrary interpretations of the Dogma. Part of the Dogma’s appeal lies not only in its generality, as seen above, but also in the fact that it makes *significant* claims about the relationship between DNA, RNA and protein. Indeed, some of the interpretations considered here regard the Dogma as significant, but each for different reasons; (as we will see) some commentators interpret the Dogma either as a thesis regarding the causal significance of DNA, as a thesis about the linear specificity of DNA, or as a thesis about the non-heritability of acquired traits. In short, many of the interpretations considered here offer reasons for the Dogma’s significance to molecular science, or biology generally. So, interpretations of the Dogma should not be restricted to the point where they end up lacking significance and are therefore arbitrary. These considerations imply that there would be no biological interest in a CD₆ that states that amino acid sequences cannot be used as predictors of the average value of Apple stock in the years 2015–2020. That thesis may be true—but it is clearly arbitrary. So, any interpretation of the Central Dogma that ends up making this an arbitrary claim like the CD₆ must be rejected. Call this the “arbitrariness condition.”

Finally, in what follows I evaluate only whether CD₁ through CD₅ accurately describe the processes underlying protein synthesis, and not whether the Central Dogma plays some other role in, or is of practical significance to, molecular science. Philosophers like Nancy Cartwright, for example, have argued that some natural laws (such as the Second Law of Thermodynamics or Newton’s Second Law of Motion), though generally false as descriptions of what actually happens, nevertheless aid in the construction of models that purport to make predictions about physical phenomena (Cartwright 1983). Others have defended other ways in which a principle may be of practical significance to science, even if it is sometimes empirically false (Kitcher 1989; Strevens 2008; Waters 2014). Some might say the same is true of the Central Dogma, i.e., that while the Dogma may be false as a description of goings-on at the molecular level, it may yet be of significance to molecular science for other reasons. However, whether this is so requires careful attention and needs to be assessed separately at another occasion. The focus here is solely on the Dogma’s descriptive accuracy, which is an important investigation in its own right.

3 Three Dogmas of Molecular Biology, Revisited

The aim of this section is to evaluate dogmas CD₁ through CD₃. Since these dogmas have figured in the literature already, the analysis here will be relatively brief. However, it is important to flag that what follows goes beyond a mere overview of the literature to date, as I will bring out several new criticisms of these three interpretations of the Central Dogma. The upshot of all of this will be further support for the contention that CD₁ through CD₃ are implausible and should be laid to rest (Griffiths and Stotz 2013; Sarkar 1996; Stotz 2006).

3.1 *The first dogma*

Recall CD₁, which says that DNA specifies RNA which specifies protein. The major challenge for this interpretation is the fact that there seem to be factors other than DNA that specify proteins. This point has been made by Karola Stotz (2006), who argues:

sequence specificity is not monopolized by DNA but is distributed among certain DNA sequences, plus regulatory RNAs, proteins, and environmental signals. If we focus on the regulation of gene expression instead of blindly taking the Dogma for granted it becomes apparent that digital and analog structures work hand in hand as they are both recruited, supported by specific environmental signals into larger multi-molecular complexes comprising DNA, RNA and proteins to synthesize gene products and regulate cellular processes. (529)

To make Stotz's point more salient, consider RNA splicing machinery as a potential source of protein specificity. RNA splicing mechanisms alter nucleobases in strands of Pre-mRNA by cutting out segments of nucleotides, and rejoining the remaining segments. Such alterations result in the expression of different proteins (Stotz 2006, 533). Given that RNA splicing mechanisms specify protein alongside DNA, it seems CD₁ fails to satisfy the generality condition specified above, and therefore must be rejected.

Alexander Rosenberg has responded to this sort of worry by noting that, while there are other causes involved in protein synthesis, DNA is the most significant cause. Specifically, Rosenberg draws on a notion that has received quite a bit of attention in the philosophy of biology literature, namely, the notion of a SAD (i.e., which is short for a causally-specific, actual, difference-maker) (Rosenberg 2006). For our current purposes, it is not necessary to give a full fledged characterization of a SAD; I do, however, return to this point in the next section. It is sufficient to note that Rosenberg's key contention is that, unlike the transcription factors sketched above, DNA makes the *biggest difference* to codons expressed in proteins—RNA editing mechanisms do not bring about such *causally-specific* differences in protein synthesis.

The major issue with Rosenberg's retort—a point which has not yet been addressed in the literature—is that he conflates CD₂ with CD₁, and thus fails to recognize that CD₂ and CD₁ are logically independent theses. As discussed, CD₂ says that DNA is the most significant factor involved in protein synthesis, while CD₁ says that DNA specifies RNA, which specifies protein. The former is a claim about *causes*, while the latter is a claim about *specificity*. Given this, the claim that genes are the most important causes involved in protein synthesis has absolutely no bearing on the claim that only DNA specifies protein. Rosenberg's counter-argument, therefore, does not reach its intended conclusion, which implies that Stotz's case against CD₁ holds.

3.2 *The second dogma*

CD₂ says that DNA is the most significant causal factor involved in protein synthesis. As noted, support has been offered for CD₂ by noting that genes constitute a SAD—i.e., a *causally-specific, actual, difference-makers*—with respect to protein synthesis (Griffiths et al. 2015; Rosenberg 2006; Stotz and Griffiths 2013; Stotz 2006; Weber 2006). An *actual difference maker* is *causally specific* if manipulating the causal variable makes for specific differences in the effect variable (Waters 2007; Rosenberg 2006). If we consider RNA synthesis as an effect, we might say that DNA is *causally specific*, because manipulating nucleotides in a strand of DNA makes for very specific differences in the RNA template. By contrast, the manipulation of RNA polymerase does not make for causally specific differences in the RNA template. In sum, a specific sort of

intervention must bring with it a specific sort of effect. This is the core of Rosenberg's appeal to the notion of a SAD to support CD₂ (Rosenberg 2006).

According to Stotz and Griffiths, the problem with this defense of the CD₂ is that there are other factors which lie beyond the cell that exhibit causal-specificity (Griffiths 2013, 87). There are "environmental signals," for instance, that make for causally-specific changes in transacting factors (Griffiths 2013, 81). In phosphorylation cascades, receptor molecules receive extracellular signals, causing Kinase molecules to activate phosphate groups within cells thereby affecting the gene-product. Since this alteration is a causally-specific one, it seems that such environmental signals satisfy the conditions for what constitute a SAD, implying that the CD₂ is also false.

3.3 *The third dogma*

In an attempt to make sense of informational construals of the Central Dogma, Sarkar introduces the notion of semiotic information (Sarkar 2004). Specifically, Sarkar uses this notion to support a construal of the Dogma which claims that DNA is the sole bearer of information with respect to protein synthesis—i.e., CD₃. Using set-theoretic concepts, Sarkar appeals to the notion of an informational relation (let's call it *i*) between sets *A* and *B* (Sarkar 2004). Set *A* contains members *a*₁ through *a*_{*n*} and all members in *A* refer to distinct sequences in strands of DNA, which correspond to codons in amino acids. For example, *a*₁ might represent the DNA sequence CTT which denotes the codon Leucine, and *a*₂ might represent the DNA sequence ACG which denotes the codon Threonine. Set *B* contains members *b*₁ through *b*_{*n*}, and all the members in *B* represent the sixty-four different codons in peptide sequences. Sarkar, then, introduces two conditions that must hold in order for it to be the case that DNA *alone* carries information:

(I1) *Differential specificity*: Suppose that *a* and *a*₁ belong to different equivalence classes of *A*. Then, if *i*(*a*, *b*) and *i*(*a*₁, *b*₁) hold, then *b* and *b*₁ must be different elements of *B*.

(I2) *Reverse differential specificity*: If *i*(*a*, *b*) and *i*(*a*₁, *b*₁) hold, and *b* and *b*₁ are different elements of *B*, then *a* and *a*₁ belong to different equivalence classes in *A*. (Sarkar 2004, 272)

To better illustrate how conditions I1 and I2 might be met, consider the DNA sequences ACG and AGG. Condition I1 is met, in part, if an informational relation exists between ACG and a particular codon. An informational relation must also exist between AGG and another codon. In this case, there is an informational relation between ACG and the codon Threonine, and also an informational relation between AGG and the codon Arginine. I1 is satisfied, since Threonine and Arginine constitute distinct elements of Set *B*. Condition I2 is met, since ACG and AGG constitute distinct equivalence classes in Set *A*. If it were the case, however, that only condition I1 was satisfied, then it could be said that *A alone* carries information for *B*, but rather that *A* carries *some* information for *B*. This is because I1 only establishes that two distinct DNA sequences code for two distinct codons in peptide sequences respectively, and not that distinct codons are the result of distinct DNA sequences.

The problem for CD₃ is the fact that protein synthesis in eukaryotic cells does not satisfy condition I2. Specifically, certain DNA sequences do not always code for, or map onto, the same codon, meaning that two distinct codons could map onto a single segment of DNA. In the case of the codon UGA in mitochondrial DNA, for example, the DNA sequence UGA codes for tryptophan as opposed to the stop codon as specified by the codon table (Sarkar 2004). Furthermore, specific DNA sequences code for different codons depending on the type of organism in

question (Sarkar 2004). This is also made salient by the familiar case of RNA editing. Such editing occurs when an RNA editing enzymes alter nucleotides in a strand of RNA, either by deleting, adding, or substituting nucleotides, thereby making significant changes in the resulting protein. RNA editing suggests that amino acids in proteins can be the result of different DNA sequences. Given this, there cannot be a strict correspondence between DNA sequences and the codons they are said to produce, which means CD₃ cannot be said to apply to all cases of protein synthesis. Since the semiotic construal of the Central Dogma fails to satisfy the generality condition, we must reject it.

4 Exploring Alternative Dogmas of Molecular Biology

CD₁ through CD₃ closely resemble Watson's interpretation of the Central Dogma, since they emphasize the role of DNA in synthesizing protein either by way of linear specificity, causal specificity, or information transfer. Given this resemblance, the proponent of the Dogma might argue that while we've managed to refute versions of Watson's dogma, we've yet to consider interpretations that align with Crick's dogma. Recall that for Crick the Central Dogma "is not the same" as the Sequence Hypothesis, "a positive statement" saying that the "(overall) transfer nucleic acid → protein did exist." Rather, the Central Dogma is "a negative statement, saying that transfers from protein did not exist" (Crick 1970). Again, the transfers Crick had in mind were those from protein to DNA, from protein to RNA, and from protein to protein (Crick 1970, 562). Let us then consider a couple negative interpretations inspired by Crick's original formulation.

4.1 *The fourth dogma*

In his now famous article "On Protein Synthesis," Francis Crick first introduces the Central Dogma as thesis about the "flow of information," which he appears to interpret as the "specification of the amino acid sequence of the protein" (Crick 1958, 144). However, if the Central Dogma is meant to be a general thesis—one that also speaks to as yet unknown transfers of information between distinct bio-molecules—this immediately leads one to wonder how this reading of the "flow of information" could apply to other biomolecules (Crick 1970, 563). On this reading, there would not be a "flow of information" from DNA to RNA (as this does not involve the specification of amino-acid sequences in a protein), and the existence of information flowing from protein to DNA would be ruled out *by fiat* (as DNA sequences are not amino-acid sequences). It is thus clear that more needs to be said about what is meant by the flow of information. (Crick himself further defines information transfer as the "directional flow of detailed, residue by residue, sequence information from one polymer molecule to another"—which, at least on the face of it, seems to beg the question [Crick 1970, 561].)

Luckily, in his original 1958 article, Crick also offers a different, broader definition of information as the "precise determination of sequence, either of bases in the nucleic acid or amino acid residues in the protein" (Crick 1958, 153). Given this broader, disjunctive characterization of the flow of information, the CD₄ can be read as denying that proteins can alter or modify other protein sequences *or* nucleobases in segments of DNA or RNA. It is this more charitable reading that will thus be at the basis of what follows—though other readings will be briefly be considered as well.

Alas, there are a number of empirical cases suggesting that the CD₄ is problematic. Let's start with the claim that proteins cannot alter or specify other proteins. Here, an immediate problem arises due to the existence of prions (Keyes 1999; Crick 1970; Prusiner 1982). Scrapie is

a protein infection (or, prion) that causes transmissible spongiform encephalopathy—or “TSE” for short—a degenerative disease causing small holes in the brain giving it a sponge-like appearance. It was discovered that the prions responsible for Scrapie were infectious, misfolded proteins that make contact with other proteins, which take on a misfolded structure also as a result of contact. Since protein infections, like Scrapie, suggest that proteins are capable of infecting other proteins, it seems we have reasons for rejecting the claim that proteins do not modify other proteins.

One may argue that prions do not *really* violate the CD₄. Specifically, it may be argued that in order for prions to serve as a true counterexample to the CD₄, prions would have to alter the amino acid sequence in a protein. Since prions alter the three-dimensional structure of the protein by infecting it with a misfolded structure, but not the primary sequence of amino acids that comprise it, the CD₄ is still true.

However, in response, it needs to be noted firstly that the nature of protein folding is still quite poorly understood and may end up being tightly linked to the protein’s amino acid sequence. Indeed, Crick himself hypothesized that protein folding “is simply a function of the order of the amino acids” (Crick 1958). If this is true, then it may very well turn out that by changing the three-dimensional structure of other proteins, the amino acid sequences of other proteins are also affected, at least in some cases. Second, even if this is not the case (though, again, this is not entirely clear), prions nevertheless affect *other* significant aspects (such as the folding) of these proteins. By disregarding these, the strict interpretation of the CD₄ thus turns out to be highly restrictive. Indeed, since these other aspects are biologically highly significant—as the case of Scrapie makes clear—a strict construal of the CD₄ ends up being quite arbitrary in a way that conflicts with the arbitrariness condition above. Hence, either way, the case of prions poses a problem for the CD₄.

When considering the other part of the CD₄—i.e., the claim that DNA/RNA-protein link cannot be reversed—the situation is quite similar. To see this, consider the processes involved in HIV transmission. HIV transmission occurs when RNA retroviruses—containing strands of RNA, the reverse transcriptase enzyme, and integrase enzyme—enter the organism’s cell by way of receptor mechanisms. Once the retroviruses make their way into the cell, a protein called “reverse transcriptase” plays an instrumental role in modifying segments of DNA, by reverse-transcribing strands of RNA to DNA. A new segment of DNA is then incorporated into the organism’s DNA template by a protein called “integrase,” causing a frameshift mutation resulting in the expression of different codons in the amino-acid sequence. Given that a series of proteins are responsible for modifying the DNA segment—reverse transcriptase is responsible for transcribing RNA into DNA, and integrase is responsible for integrating the segment into the existing DNA—this suggests that proteins *can* specify DNA.

Now, one might contest this case by pointing out that while (a) reverse transcriptase reverse transcribes RNA to DNA, and (b) integrase incorporates the new strand of DNA into the existing strand of DNA, this does not amount to amino acids being “reverse translated” and transcribed into DNA. One might therefore conclude that the CD₄ is safe after all.

However, it is again not clear that this is true. This is because, most straightforwardly, the CD₄ says that proteins cannot *determine* DNA. As such, this “*determination*” says nothing about the need for reverse translation: if protein *P* (using an RNA intermediate) can change the nucleotide sequence in strand of DNA *S*, then this appears to be a case of *P* determining *S*, independently of whether *P* was strictly reverse translated into *S*. Given this, the case of HIV transmission clearly violates the dogma. Reverse transcriptase and integrase play a significant role in determining the nucleic acids in DNA segments, as discussed above.

At any rate, if a stricter interpretation of the CD₄ is adopted, according to which a protein would need to be reverse translated into DNA to pose a counterexample to it, the CD₄ again ends up looking arbitrarily restricted. What cases of retroviruses show is that parts of the DNA-RNA-protein determination *can* be reversed: this is not a fully one-way affair. This is biologically significant. By contrast, the strict, reverse-translation-based interpretation of the CD₄ appears to a somewhat arbitrary claim in line with the fictitious CD₆ above. Again, the CD₄ fails—either because it is false or because it falls foul of the arbitrariness condition.

At this point, a final objection to these cases needs to be addressed. In particular, one may be concerned that diseases like Scrapie and HIV are sufficiently rare to be able to be accommodated as outliers by the CD₄. Crick (1970) hints at this when he says:

we know enough to say that a non-trivial example showing that the classification was wrong could be an important discovery Any of these would be of the greatest interest, but they could be accommodated into our thinking without undue strain. (562–563)

However, as noted above, this response will not work, as they undermine the generality constraint in interpretations of the Central Dogma. What makes the Central Dogma a “dogma” is precisely the fact that it describes general facts about all cases of protein synthesis (and bio-molecular interactions, generally). Furthermore, what makes the Central Dogma interesting as principle in biology is its ability to describe a general process at the bio-molecular level. We should therefore expect any construal of the Dogma to apply in all cases, and not just some. On top of this, it is anyway the case that scrapie (for example) is considered part of a *family* of prion diseases that are actually quite common in the animal kingdom (Keyes 1999). In fact, specific prions have been found to be of benefit to the organism. So, it is not true that such cases should be dismissed as outliers. (The same goes for the proteins responsible for reverse-transcribing RNA to DNA, as the case of HIV transmission makes clear.) In short: CD₄, too, is either false or arbitrarily restricted.

4.2 *The fifth dogma*

Before we consider the fifth and final interpretation of the Central Dogma, I should note that for many commentators the Central Dogma isn't merely a thesis characterizing a range of possible bio-molecular interactions, but also a thesis about the evolution and development of an organism's phenotypic traits. Specifically, other than Dawkins (1982), many authors interpret the Dogma as being a thesis about the heritability of acquired traits.

Most widely, the central dogma was the restatement—radical, absolute—of the reason why characteristics acquired by an organism in its life but not from its genes cannot be inherited by its offspring. “Once information has passed into protein *it cannot get out again*.” (Judson 1979)

The Central Dogma is ... the final nail in the coffin of the inheritance of acquired characteristics. (Wilkins 2002)

There were two aspects of Crick's lecture that related to evolutionary thinking. The first was that the Central Dogma supported the neo-Darwinian view that it was impossible for any character that was acquired during an organism's life to affect its hereditary characters. This provided support for the widespread hostility to the view that had been held by Darwin, Lamarck, and others, according to which, patterns

of use and disuse could lead to changes in the frequency of characters in subsequent generations. (Cobb 2017)

These are just a few examples where commentators have sketched a tight-knit connection between the Central Dogma and the claim that heritable traits cannot be acquired by an organism.

Given this connection, the proponent of the Dogma might argue that the interpretations above (CD₁ through CD₄), while significant in that they depict goings on at the molecular level, nevertheless gloss over what the Central Dogma is *really* trying to say. The proponent of the Dogma might say that we've failed to consider interpretations according to which it is impossible for organisms to acquire non-culturally heritable traits by way of the environment. The restriction to *non-culturally* heritable traits is important, as no one seems to question that non-genetic—e.g., cultural learning—inheritance can be environmentally-driven. The question, though, is not whether such traits can be inherited non-genetically via learning, but rather whether it is possible for acquired traits to be genetically inherited. To this, a defender of CD₅ says: no.

(Before we delve into CD₅, a quick note regarding this interpretation is in order. When Lamarck himself discussed the inheritance of acquired traits, he specifically referred to events in which an organism's traits are modified through habits of use and disuse. For Lamarck, such modifications are somehow conserved within the organism, and then passed on to the organism's offspring. I will not be considering Lamarck's interpretation here. Instead, I will opt for a formulation that best comports with our contemporary understanding of genetic expression.)

CD₅ holds that non-culturally heritable traits proceed only from mutations in the organism's genes and no place else. CD₅, therefore, denies that environmental factors affect the organism in a way that produces non-culturally heritable phenotypic traits. To better illustrate CD₅, consider the phenotypic trait of having blue eyes. Suppose, for argument's sake, that this phenotypic trait *P* surfaces in a subset *A* of organisms in population *B*, but not as a result of genetic mutation. Rather, suppose *P* was acquired by *A* as a result of a specific nutrient. According to CD₅, *P* will not be non-culturally passed onto *A*'s offspring, because the trait itself did not emerge as a result of mutation. In short, just as bruises are not acquired and passed on to an organism's offspring, eye color cannot be acquired environmentally and non-culturally passed onto an organism's offspring.

To see why CD₅ fails, consider the phenomenon of RNA interference (or RNAi). RNAi occurs when strands of RNA inhibit the expression of genes. Vastenhouw et al. demonstrated that such effects result in the temporary heritable transmission of traits (Vastenhouw et al. 2006). Specifically, Vastenhouw et al. demonstrated that by injecting double-stranded RNA into the *ceb-13* gene in nematodes, one could silence the *ceb-13* in the organism, an effect which persisted for up to 2–3 generations. These findings suggest that external factors do in fact affect an organism's DNA, giving rise to non-culturally heritable traits, which means that CD₅ is false.

Note again that it is not possible to rescue the Dogma by arguing that the exception sketched above, rather than undermining CD₅, proves the rule. So, while it may be that cases of RNA interference are quite short-lived and rare, they still exist, and violate the CD₅. Given the generality constraint of interpretations of the Dogma, the CD₅ must also be rejected.

5 Conclusion

In this paper, I've shown that if the Central Dogma is sketched as a thesis about specificity—whether linear or causal—then we must acknowledge the numerous factors involved in protein

specificity. If construed as a thesis about semiotic information, then we must reckon with factors disrupting mappings between nucleic acids in DNA segments and amino acids in proteins. Even if we depart from contemporary expressions of the Dogma, and attempt to sketch it as a thesis resembling Crick's original formulation, there are empirical cases which render broader interpretations of the Dogma false. If conceived as a negative thesis about what protein can't specify, we run into problems due to the existence of protein infections like scrapie (which suggest that proteins *do* specify proteins) and the existence of RNA retroviruses (which suggest that proteins *do* specify DNA). Finally, if construed as a thesis about the denial of the heritability of traits, then we must come to terms with the effects of RNA interference. In conclusion, the Central Dogma is empirically inadequate no matter how we slice it.

Acknowledgments

Thanks to Armin Schulz, John Symons, Sarah Robins, Corey Maley, and David Tamez for helpful comments on earlier drafts of this manuscript.

Literature cited

- Alberti, Simon, Randal Halfmann, Oliver King, Atul Kapila, and Susan Lindquist. 2009. "A Systematic Survey Identifies Prions and Illuminates Sequence Features of Prionogenic Proteins." *Cell* 137 (1): 146–158. doi:10.1016/j.cell.2009.02.044.
- Burkhardt, Richard W. 2013. "Lamarck, Evolution, and the Inheritance of Acquired Characters." *Genetics* 194 (4): 793–805. doi:10.1534/genetics.113.151852.
- Boyd, Robert, and Peter J. Richerson. 1996. "Why Culture is Common, But Cultural Evolution is Rare." *Proceedings of the British Academy* 88: 77–94.
- Cartwright, Nancy. 1983. *How the Laws of Physics Lie*. Oxford: Clarendon Press.
- Cartwright, Nancy. 1999. *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.
- Cartwright, Nancy, Towfic Shomar, and Mauricio Suárez. 1995. "The Tool Box of Science: Tools for the Building of Models with a Superconductivity Example." *Poznan Studies in the Philosophy of the Sciences and the Humanities* 44: 1 37–149.
- Crick, Francis. 1958. "On Protein Synthesis." *Symp. Soc. Exp. Biol.* 12: 138–63.
- Crick, Francis. 1970. "Central Dogma of Molecular Biology." *Nature* 227 (5258): 561–563. doi:10.1038_227561a0.
- Crippen, Gordon M. 1991. "Prediction of Protein Folding From Amino Acid Sequence Over Discrete Conformation Spaces." *Biochemistry* 30 (17): 4232–4237. doi:10.1021/bi00231a018.
- Cobb, Matthew. 2017. "60 Years Ago, Francis Crick Changed the Logic of Biology." *PLOS Biology* 15 (9): e2003243. doi:10.1371/journal.pbio.2003243.
- Dawkins, Richard. 1982. *The Extended Phenotype*. Oxford: Oxford University Press.
- De Tiège, Alexis, Koen Tanghe, Johan Braeckman, and Yves Van de Peer. 2014. "From DNA-to NA-Centrism and the Conditions for Gene-Centrism Revisited." *Biology & Philosophy* 29 (1): 55–69. doi:10.1007/s10539-013-9393-z.
- Fisher, Susie. 2010. "Not Beyond Reasonable Doubt: Howard Temin's Provirus Hypothesis Revisited." *Journal of the History of Biology* 43 (4): 661–696. doi:10.1007/s10739-009-9202-3.

- Godfrey-Smith, Peter, and Kim Sterelny. 2016. "Biological Information." In *Stanford Encyclopedia of Philosophy* (Summer 2016 Edition), edited by Edward N. Zalta. Stanford University. <https://plato.stanford.edu/archives/sum2016/entries/information-biological/>.
- Graur, Dan. 2018. "The Fallacious Commingling of Two Unrelated Hypotheses: 'The Central Dogma' and 'DNA Makes RNA Makes Protein.'" *Judge Starling*. <http://judgestarling.tumblr.com/post/177554581856/the-fallacious-commingling-of-two-unrelated>.
- Griffiths, Paul, and Karola Stotz. 2013. *Genetics and Philosophy: An Introduction*. Cambridge: Cambridge University Press.
- Griffiths, Paul E., Arnaud Pocheville, Brett Calcott, Karola Stotz, Hyunju Kim, and Rob Knight. 2015. "Measuring Causal Specificity." *Philosophy of Science* 82 (4): 529–555. doi:10.1086/682914.
- Judson, Horace F. 1979. *The Eighth Day of Creation*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
- Keller, Evelyn F. 2009. *The Century of the Gene*. Boston: Harvard University Press.
- Keller, Evelyn F. 1995. *Refiguring Life: Metaphors of Twentieth-Century Biology*. New York: Columbia University Press.
- Keyes, Martha E. 1999. "The Prion Challenge to the 'Central Dogma' of Molecular Biology, 1965–1991: Part I: Prelude to Prions." *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 30 (1): 1–19. doi:10.1016/S1369-8486(98)00028-4.
- Lewis, David. 2000. "Causation as Influence." *The Journal of Philosophy* 97 (4): 182–197. doi:10.2307/2678389.
- Smith, John M. 2000. "The Concept of Information in Biology." *Philosophy of Science* 67 (2): 177–194.
- Mallela, Arka, and Kazuko Nishikura. 2012. "A-to-I Editing of Protein Coding and Non-coding RNAs." *Critical Reviews in Biochemistry and Molecular Biology* 47 (6): 493–501. doi:10.3109/10409238.2012.714350.
- Wilkins, Adam S. 2002. "Interview with Ernst Mayr." *BioEssays* 24: 960–973. doi:10.1002/bies.10167.
- Mizutani, Satoshi, and Howard M. Temin. 1970. "An RNA-Dependent DNA Polymerase in Virions of Rous Sarcoma Virus." *Nature* 226 (5252): 1211–1213. <https://www.ncbi.nlm.nih.gov/pubmed/4316301>.
- Morange, Michel. 2008. "What history Tells Us XIII. Fifty Years of the Central Dogma." *Journal of Biosciences* 33 (2): 171–175. doi:10.1007/s12038-008-0034-7.
- Morange, Michel. 2009. "The Central Dogma of Molecular Biology." *Resonance* 14 (3): 236–247. doi:10.1007/s12045-009-0024-6.
- Morrier, Douglas. 2019. "Reverse Transcriptase". In *Encyclopedia Britannica*. <https://www.britannica.com/>.
- Morris, James. 2013. *Biology: How Life Works*. London: Macmillan Higher Education.
- Oyama, Susan. 2000. *The Ontogeny of Information: Developmental Systems and Evolution*. Durham: Duke University Press.
- Oyama, Susan. 2000. *Evolution's Eye: A Systems View of the Biology-Culture Divide*. Durham: Duke University Press.
- Pearl, Judea. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.

- Prusiner, Stanley B. 1982. "Novel Proteinaceous Infectious Particles Cause Scrapie." *Science* 216 (4542): 136–144. doi:10.1126/science.6801762.
- Prusiner, Stanley B. 2013. "Biology and Genetics of Prions Causing Neurodegeneration." *Annual Review of Genetics* 47: 601–623. doi:10.1146/annurevgenet-110711-155524.
- Psillos, Stathis. 2002. *Causation and Explanation*. Montreal: McGill-Queen's University Press.
- Raineri, Deanna. 2001. *11th Hour: Introduction to Molecular Biology*. Hoboken: Wiley.
- Rosenberg, Alex. 2006. "Is Epigenetic Inheritance a Counterexample to the Central Dogma?." *History and Philosophy of the Life Sciences* 28 (4): 549–565. <https://www.ncbi.nlm.nih.gov/pubmed/18351052>.
- Rosenberg, Alexander. 2008. *Darwinian Reductionism: Or, How to Stop Worrying and Love Molecular Biology*. Chicago: University of Chicago Press.
- Sarkar, Sahotra. 2004. *Molecular Models of Life: Philosophical Papers on Molecular Biology*. Cambridge: MIT Press.
- Sarkar, Sahotra. 1996. "Biological Information: A Skeptical Look at Some Central Dogmas of Molecular Biology." *Boston Studies in the Philosophy of Science* 183: 187–232.
- Schrödinger, Erwin. 1992. *What is life?: With Mind and Matter and Autobiographical Sketches*. Cambridge: Cambridge University Press.
- Shah, Neel H., and Tom W. Muir. 2014. "Inteins: Nature's Gift to Protein Chemists." *Chemical Science* 5 (2): 446–461. doi:10.1039/C3SC52951G.
- Shafer, Robert W., Kathryn Dupnik, Mark A. Winters, and Susan H. Eshleman. 2001. "A Guide to HIV-1 Reverse Transcriptase and Protease Sequencing for Drug Resistance Studies." *HIV Sequence Compendium* 1–51. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3274565/>.
- Smith, John M. 1993. *The Theory of Evolution*. Cambridge: Cambridge University Press.
- Sturtevant, Alfred H. 2001. *A History of Genetics*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
- Stotz, Karola. 2006. "Molecular Epigenesis: Distributed Specificity as a Break in the Central Dogma." *History and Philosophy of the Life Sciences* 28 (4): 533–548. <https://www.ncbi.nlm.nih.gov/pubmed/18351051>.
- Šustar, Predrag. 2007. "Crick's Notion of Genetic Information and the 'Central Dogma' of Molecular Biology." *The British Journal for the Philosophy of Science* 58 (1): 13–24. doi:10.1093/bjps/axl018.
- Strasser, Bruno J. 2006. "A World in One Dimension: Linus Pauling, Francis Crick and the Central Dogma of Molecular Biology." *History and Philosophy of the Life Sciences* 28 (4): 491–512. <https://www.ncbi.nlm.nih.gov/pubmed/18351048>.
- Starokadomskyy, P. L. 2007. "Protein Splicing." *Molecular Biology* 41 (2): 278–293. doi:10.1134/S0026893307020094.
- Temin, Howard M., and David Baltimore. 1972. "RNA-directed DNA Synthesis and RNA Tumor Viruses." *Advances in Virus Research* 17: 129–186. doi:10.1016/S0065-3527(08)60749-6.
- Thieffry, Denis, and Sahotra Sarkar. 1998. "Forty Years Under the Central Dogma." *Trends in Biochemical Sciences* 23 (8): 312–316. doi:10.1016/S0968-0004(98)01244-4.
- von Wright, Georg H. 2004. *Explanation and Understanding*. Ithaca: Cornell University Press.
- Waters, C. K. 2004. "What Was Classical Genetics?." *Studies in History and Philosophy of Science Part A* 35 (4): 783–809. doi:10.1016/j.shpsa.2004.03.018.
- Waters, C. K. 2007. "Causes that Make a Difference." *The Journal of Philosophy* 104 (11): 551–579.

- Watson, James D. 1965. *Molecular Biology of the Gene*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
- Weber, Marcel. 2006. "The Central Dogma as a Thesis of Causal Specificity." *History and Philosophy of the Life Sciences* 28 (4): 595–609.
- Woodward, James. 2005. *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.

© 2019 Author(s)

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits anyone to download, copy, distribute, display, or adapt the text without asking for permission, provided that the creator(s) are given full credit.

ISSN 2475-3025