
THE ADVENTURES OF INFORMATION IN THE LAND OF HI-FI AND HD

JUAN FELIPE GUEVARA-ARISTIZABAL

ABSTRACT. The concept of information has acquired a substantial power in biology, becoming the mayor metaphor of modern biology. This has allowed the development of disciplines such as recombinant DNA technology, transgenic organisms, genome sequencing, and what we now call biotechnology. Yet its birth was much humbler and did not encompass so many options. The concept of information in biology could achieve a much-refined stance than just the equivalent of genetic information by enlarging its conceptual scope, not limited to molecular terms but dealing with higher levels of organization. It is intended here to lay the bases for an updated comprehension of information, widening its perspective within this land of hi-fi and HD that biology has become.

KEY WORDS. Information, genetic program, digital information, positional information, structural information, complementarity, biosemantics.

1. INTRODUCTION

Information is one of many concepts borrowed from non-biological disciplines that certainly propelled the advance of biological sciences towards new frontiers. The history of the last sixty years of biology is unconceivable without this concept and, perhaps, its influence will last even longer. No matter what biological discipline or theory you are looking from, the notion of information will always get in the way. Understandably, some disciplines have been more influenced by it than others, for example, genetics, molecular biology, and developmental biology (Maynard Smith 2000; Godfrey-Smith 2007). If the concept of information would not had taken its part in biology, no such advancements as recombinant DNA technology, transgenic organisms, gene therapy, or synthetic biology would have occurred. A good corollary for this type of research would be the recent works of the J. Craig Venter Institute that have rendered as historical results the generation of the first synthetic genome (Smith, et al.

Instituto de Biología, Universidad de Antioquia, Medellín, Colombia. / juanfga@gmail.com

2003) and the chemical synthesis of a bacterial chromosome and its assembly into a chromosome-less bacterial cell to ‘create’ the first synthetic cell (Gibson, et al. 2010), as they claim in the title of their publication, when it should actually be considered as the first synthetic genome to be successfully replicated within a cell. Plus, the history of this kind of scientific applications shows its implications within the social and ethical dimension of science. A good example of these issues is the 1975 Asilomar Conference on recombinant DNA technology organized by Paul Berg. Its main focus was the potentially harmful products of this technology for human health and the whole environment, due to new virus or bacterial strains that could escape from labs and contaminate the milieu easily (Berg, et al. 1975). Twenty years later, things had changed a lot. None of the dangers assessed by the Asilomar Conference became an actual threat for humanity or the environment; practically, these fears became real only in films and videogames. Nonetheless, another kind of problems originated—the life that was manipulated in labs was starting to be patented in courts (Berg and Singer 1995). Thanks to the concept of information, biological sciences were able to enter the larger game of intellectual property and patents, and labs found another way of increasing their incomes. This link would take us directly to another discussion on bioethics; yet in spite of how brilliantly seducing it could be, it is not at the scope of this essay. Anyhow, the authors of both papers focused their efforts on highlighting the beneficial effects of recombinant DNA technology despite the controversies inside and outside academic circles. Hence, the questions that really pertain this essay remain to be examined: How did the concept of information entered biological sciences? And, what exactly does it mean in a biological theoretical framework?

2. WHEN A METAPHOR BECOMES THE BEST ALLY OF SCIENCE

It is difficult to establish a precise date for the entrance of information in the biological scenario. The first mention of information in biology seems to have been done by Weismann at the end of the nineteenth century, when he compared the process of inheritance with a specific information-transducing channel—the telegram (Maynard Smith 2000). There was no formal information theory by that time. Its bases were only laid by the 1920s, and by the 1940s it already counted with a mathematical theory describing information transmission, i.e., a theory of communication, thanks to the works of Claude Shannon and Norbert Wiener.

On the side of biological sciences, when information intertwined with biology, the particles of inheritance were already identified as genes made of DNA, the double-helical structure of DNA was already proposed and its replication mechanism elucidated. Perhaps the first explicit mention was

made by Francis Crick in his seminal 1958 paper, where he spoke of 'information flux' from nucleic acids to proteins and, more importantly, of the coding nature of the DNA sequence of nucleotides so important for the process of protein synthesis, which was refined and postulated in the central dogma of molecular biology (Crick 1970). It was a concept borrowed from another discipline, introduced as a novel metaphor to construct, linguistically, an explanation for genetic and molecular data (Fox Keller 2003)—trying to use the familiar to explain the unfamiliar (Maynard Smith 2000). The notion of 'coding' introduced by Crick turned the information contained in DNA comparable to the information in Shannon's theory as describing the transmission of a message through a channel. In addition, one of the most important contributions in Shannon's theory was the possibility to measure information in terms of bits, binary digital units, 1s or 0s. It allowed the introduction of another metaphor into biology, this time not from information theory but from computer science¹—the basic layout of a program is always an algorithm, which is then translated into the language of bits in order for the program to run on a computer. This notion of program is very appealing to the phenomena observed and described in genetic and molecular terms. Some biologists noted this fact, and in 1961 the concept of 'genetic program' appeared for the first time in scientific publications.

The interesting fact about that first social presentation of the genetic program concept in biology is that it was made by two very different kinds of biologists. On the one hand, the Jacob-Monod team in a publication in the *Journal of Molecular Biology*. On the other hand, Mayr in a paper published in *Science*. Jacob and Monod published in June, while Mayr did it on November; the operon team took the lead for scant three months. Also, their formulation of the concept is by far clearer and closer to what the metaphor of the computer program tried to achieve. However, I would like to assess Mayr's concept first.

The minimal statement of 'program' in Mayr (1961) describes it as purely mechanistic purposiveness. It might sound contradictory to include on the same sentence, and referring to the same term, the words mechanistic and purpose; at least in the context of the physicalist traditions of philosophy of science, where mechanistic explanations of the world had no space for purpose. Yet, in order to achieve another counter-argument against the use of the traditional philosophy of science in biology, Mayr (1992) combines both in the concept of program as it gives a new complexion to the problem of goal-directedness, one that suits biology perfectly. The detailed and extended definition of program appeared much later, during the last days of Mayr: coded or prearranged information that controls a process (or behavior) leading toward a given end or goal (Mayr 2004). A program implies two other characteristics: it is

something material and it exists prior to the initiation of the process it codes for. As material entities, programs lie on molecules and other organic components, so its existence does not suppose any conflict with natural laws (Mayr 2004). Another important feature of programs is that they contain not only the blueprint but also the instructions of how to use the information of such blueprint; the key to understand such feature is that the blueprint and its instructions are the result of the same process.

Each particular program is the result of natural selection and is constantly adjusted by the selective value of the achieved endpoint ² (Mayr 1974), therefore, programs are also the result of evolution. However, Mayr (1982) posits that it is the endpoint that produces the selective pressure, which causes the historical construction of the genetic program. Here we find a flagrant contradiction in terms of the concepts of evolutionary biology. For any feature that evolves by natural selection, the selective value of a feature refers to the increase or decrease in fitness that confers to organisms due to their interaction with the environment, with external factors, whether living or non-living, that exert a selective pressure over the feature. It is clear, then, that the selective value is a dynamic property of a feature and its interaction with the environment, while the selective pressure is a property of the environment. They are like two different coins, not the two faces of the same one. As a consequence, it is conceptually impossible to conceive how a feature could be, at the same time, the determinant of the selective value and of the selective pressure. That is the dilemma behind Mayr statements on the endpoint: in 1974 he speaks of it as the responsible of the selective value of a program and, then, in 1982, it is responsible for the selective pressure acting on the program. During this eight-year gap, we could think that Mayr changed his views on the role of the endpoint. Anyway, when one reads Mayr writings it is easy to see that he is not one of those persons that modify his ideas very easily; he might complement most of them, but his basic positions and reflections up to the 1970s did not change in his subsequent works. It is my opinion, then, that this conceptual misdemeanor might not be the result of a careful reflection on the topic that took him to change his ideas. Perhaps, it is more suitable to think that this kind of contradiction results from the nature of the subject itself: endpoints still have a metaphysical halo that is hard for a scientist to take off, and any attempt to do so from the side of natural sciences implies some risks that even a biologist like Mayr could not escape from.

Now we turn the page to take a look at the Jacob-Monod conceptualization of genetic program. In order to understand their ideas we need to take a look into how the effects of genes in organisms were conceived. The first approach can be found in the notion of 'gene action' (Fox Keller 2003), which describes the direct involvement of genes in protein synthesis, and

it is closely related to the one gene-one enzyme hypothesis. As a consequence, it deals exclusively with structural genes, where no notion of regulation is included. The problem that many biologists found in this interpretation of genetic phenomena was that it could not explain the process of cell differentiation in pluricellular organisms—how does any pair of cells, bearing the same genetic material, could specialize and display strikingly dissimilar morphologies and produce distinct physiological actions? This question haunted biologists until research on a unicellular organism shed light on the regulatory mechanisms involved in the differentiation process observed during development of pluricellular ones. The work of Jacob and Monod on *E. coli* and their operon theory fulfilled the missing spot highlighted by the question on development, and it also left the road free for the advent of the concept of genetic program (Morange 2005). Let me quote them to better understand the point mentioned:

According to the strictly structural concept, the genome is considered as a mosaic of independent molecular blue-prints for the building of individual cellular constituents. In the execution of these plans, however, co-ordination is evidently of absolute survival value. The discovery of regulator and operator genes, and of repressive regulation of the activity of structural genes, reveals that the genome contains not only a series of blue-prints, but a coordinated program of protein synthesis and the means of controlling its execution (Jacob and Monod, 1961, p. 354).

The genetic regulatory mechanisms they found to work on *E. coli* were used as the foundations for a more general theory regarding the process of differential gene activation during development—recall the quote “anything found to be true of *E. coli* must also be true of elephants” (Monod and Jacob 1961, p. 393). As a consequence, the notion of program “suggests a plan of procedure, a schedule, or even a set of instructions” (Fox Keller 2003, p. 136). This is very important in the comparison to a computer program, because such instructive nature emphasizes the similarity to the algorithm underlying the program.

In the history of biology it is not surprising to find such differences as between Mayr’s conceptions and those of Jacob and Monod. The genetic program concept is no exception, and the distinction can be easily grasped from the paragraph quoted above. Mayr (1974, 1992, 2004) defines the genetic program in terms of the blueprint as well as the instructions. What is then the difference between genetic program and genome? There would be none since the genome also contains the blueprints for protein synthesis and the set of instructions on how and when to synthesize them³. Consequently, there would be no need to allege for another concept: just stay with the genome. In contrast, Jacob and Monod consider the blue-

prints and the instructions, the genetic program, as two different parts that in sum make up the genome. Considering the genetic program ⁴ as the instructions makes more sense to the metaphor of a computer program since an algorithm is at its base. What the algorithm describes is a series of decision nodes that guides a certain process depending on the conditions or requirements imposed at each node (Trevors and Abel 2004). Yet the algorithm is an abstract construction. Therefore, genetic programs are algorithms instantiated, i.e., recorded into a physicality of choices, into the physical medium of nucleotide sequences (Trevors and Abel 2004). In conclusion, what the algorithm deals with is a set of instructions on what to do depending on the conditions achieved at each decision node; it does not contain any blueprint inasmuch an algorithm can produce many different outcomes depending on the pathway taken at each node. Maynard Smith (2000) and Godfrey-Smith (2007) take the argument further and declare that the genome is not a description of the adult form, but a recipe of how to make it. As a consequence, there is no place for blueprints in the genome—it is only instructions.

Nonetheless, the metaphor of the program is also completed by the fact that the genetic program is also written in a digital language. Even more, not only the program is written in such language, but also the blueprints; in other words, the whole genome is digital. This is a key point connecting also genetic programs and DNA with the technical definitions of Shannon's theory of information. Additionally, computer programs had a very specific meaning encoded in the linear/digital sequence of bits (Fox Keller 2003), which adds a semantic dimension to the information stored in the DNA. As a result, information, the program, or DNA can be studied from two different standpoints: as the linear sequence of nucleotides (comparable to bits, although, instead of two bits, DNA has four) and its transmission, or as the meaning encoded in that linear sequence. Under these circumstances, the concept of information pervaded all of biology, to the point that Jacob (1970) declared that the biologist actually deals with the evolution of information. Therefore, what exactly information means in a biological context?

3. TAKING THE METAPHOR DOWN TO EARTH

The use of the information metaphor in biological sciences ranges from the term itself to a whole set of terms derived from information theory that have become regular for any biologist's argot. It includes terms like 'transcription', 'translation', 'code', 'redundancy', 'synonymous', 'messenger', 'editing', 'proofreading', and 'library', among others (Maynard Smith 2000). Despite its widespread usage, the concept of information derived from Shannon is not the same that biologists actually refer to. Nowadays, there are several concepts of information—at least three can be outlined so far

in this text: statistical information, physical information, and semantic information (Harms 2006). The first one, statistical information, or Shannon information, is a quantitative concept measured in bits (binary digits) and concerned with the transmission of a certain message, which involves an information source, a transmitter, a channel, a receiver, and a destination⁵ (Shannon 1948). Within this frame, anything is a source of information if it has a number of alternative states that might be recognized on a specific occasion, and any variable carries information about the source if its state is correlated with that of the source (Godfrey-Smith 2007). What makes Shannon's theory statistical is the fact that information arises from the improbability or uncertainty of a state—the less common a state is, the more information is generated by the system being in that state (Harms 2004, 2006). The statistical concept is relevant when dealing with replication, transcription, translation and genetic inheritance—processes concerned with the transmission of genetic information, whether from DNA to RNA or protein, or from one generation to the next. It is by characterizing information in terms of improbability and uncertainty that takes us to the second concept.

Physical information is based on the long-standing analogy between Shannon's information and entropy⁶ (Boltzmann's statistical interpretation). If entropy characterizes the amount of uncertainty or 'unknowledge,' then any decrease of uncertainty is equivalent to an increase of information (Eigen and Schuster 1977). Eigen and Schuster's formulation might seem contradictory with the previous statement: if the less common a state is (the more uncertainty there is about it), the more information generated by the system, how can a decrease in uncertainty be equivalent to an increase in information? Since the relationship between entropy and information has been stressed countless times in living systems, let me answer this question using a eukaryotic cell as our study system. In its most basic interpretation, entropy denotes the number of ways in which a system may be arranged (Baierlein 2003), i.e., the disposition of the molecules forming it. A system has the most entropy when the probability of 'locating' a given molecule in the system is almost the same for every place within the system—when the system is highly disordered. On the contrary, when the probability of finding a molecule in a specific place is very high, then the system is ordered and the entropy is small. This happens in the cell. For example, the probability of finding DNA in the cytoplasm of a eukaryotic cell is almost zero, unless there is some damage in the nuclear, mitochondrial or plastid membranes. Thus, what the eukaryotic cell does is to compartmentalize most of its contents, so they will be found at very specific places. As a consequence, entropy is strongly reduced. This ordered form is one of the many states the system called 'eukaryotic cell' could exhibit, and it is the most improbable state considering the second

law of thermodynamics—a system tends to the highest entropy state, and that is exactly what the laws of physics and chemistry try to do with the cell. Yet, the cell actively consumes energy to keep its order, to keep itself in the most improbable state. Therefore, the cell generates the more information—although the uncertainty of the components of the system (the cell) decreases, the uncertainty of the state of the system as a whole resulting from the organization of those components is extremely high.

The DNA, the coding material by definition, contained in the eukaryotic cell model previously assessed serves as the link to the last concept of information. The semantic concept deals with the meaning and the convention rules that allow different entities to understand the meaning of the message (Harms 2006; Godfrey-Smith 2007). Even so, the semantic interpretation should not be confused with its pragmatic interpretation, the response elicited in the destination. As we will see in further paragraphs, the semantic approach is not clear enough on whether biological information necessarily has a coding nature or depends entirely on the set of rules used by the destination to grasp and decode the message. Thus, this latter interpretation clearly shows the pragmatic side of information.

Which of these concepts is pertinent to biological sciences? All of them have something to do with the discipline or theory of biological sciences. Many have argued that the statistical concept has little if nothing to do in biology (Maynard Smith 2000; Godfrey-Smith 2007) based on the following statement by Shannon (1948):

The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one selected from a set of possible messages (p. 379).

The problem that arises from this quote is that they consider that the actual problem of information in the realm of biology is that of meaning, semantics, only because of the coding nature of DNA. However, they are forgetting that DNA information is constantly transmitted, either during gene expression or inheritance, and they might be also omitting the fact that there is a recent branch of biology in which Shannon's concept of information fits perfectly: bioinformatics. In bioinformatics, sequences of DNA are studied and compared to find any similarity between them regardless of their meaning, i.e., what they code for. It is closer to a vision of information centered on syntax—it deals with the molecular structure and the capacity of biological information carriers (the exact sequence of bases that build the DNA molecule) (Schuster 2002), which is more related to

Shannon's theory than to semantics. Also, whenever we talk about mutations we are indirectly taking into account the statistical concept, since we are dealing with the result of an unfaithful process of information transmission. The bias is that every time we discuss mutations we are only interested in their impact on what the mutated sequence codes for, which is its semantic repercussion. A clear and outstanding example of the importance of Shannon's information theory in biology is the neutral theory of evolution, worked out by Kimura (1968), thanks to its mathematical treatment built around mutation rates and the neutrality of most mutations—a theory that has more to do with how faithfully information is transmitted during inheritance rather than how the specific changes modify the code and its products. In the case of the physical concept, it is also relevant when considering any biological entity as a physical system far from equilibrium. Irrespective of the conceptual diversity, mainstream biologists and philosophers of biology consider semantics the central problem of biological information (Millikan 1989; Maynard-Smith 2000; Jablonka 2002; Schuster 2002; Godfrey-Smith 2007). Therefore, the rest of this section will deal with this notion.

Since the concept of program was so important and the molecular structure of DNA was so appealing to the syntax described by communication theory, the problem of semantics focused on the coding properties of DNA, its inheritance and processing. Genes contain information about the proteins they make, as well as whole-organism phenotypes (Godfrey-Smith 2007). However, we cannot forget that proteins are not the only genic products. There are some genes that are only transcribed but not translated, as in the case of transfer RNA (tRNA), ribosomal RNA (rRNA), and micro-RNAs which are part of the system of RNA interference (RNAi), all of which are fundamental for the appropriate translation of a messenger RNA (mRNA): tRNAs are the connector molecules that allow the codons in the mRNA to guide the assemblage of amino acids (Lilley 2003), rRNAs are the basic constituents of the ribosome and they direct the formation of the peptidic bond between amino acids (Noller 2005), and RNAi can inhibit the process of translation by binding to specific mRNAs or even genes, becoming a novel mechanism of gene silencing (Zamore 2006). Having made such conceptual clarification, let us continue with Godfrey-Smith's claim. The first part is easy to grasp: a gene is transcribed into an mRNA sequence (or it could be already an RNA molecule, as in some viruses) to be later translated into a protein depending on the sequence of codons (three-letter-RNA words that specify for a given amino acid) that direct the assembling of amino acids in ribosomes. Proteins are coded in genes, while codons encode for amino acids; therefore, DNA contains a code that must be decoded in order to obtain its various products, another molecule that can perform a certain function within the cell, whether an RNA or a protein

(in this case, only translation is examined). This interpretation is directly related to the problem of information flux or transmission and can be summarized in figure 1. However, this scheme lacks a suitable equivalent in the DNA context for the place occupied by the coder in the Morse code case. We find that both forms of transmission share a source (words or DNA), a message (Morse code or mRNA), a decoder (decoder or ribosome), and a product (words or protein 7). In this case, Maynard Smith compares the process of communication through the Morse code with that of translation in a cell. Therefore, it is understandable that the source and the product are different in the process of translation.

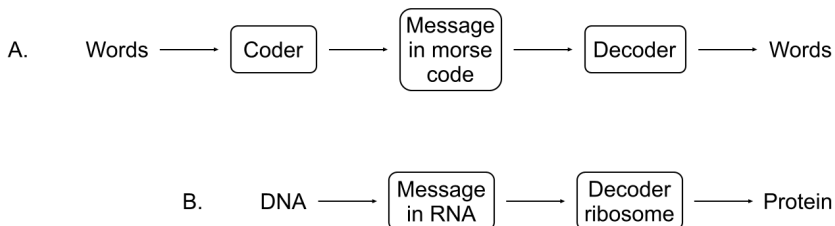


FIGURE 1

Comparison of A, human message transmission in Morse code, and B, translation of the message coded in DNA into protein (based on Maynard Smith 2000).

That would not be the case in the second kind of information contained in the genes according to Godfrey-Smith: a whole-organism phenotype. The particularity of this case is that we are no longer dealing with transmission of information in a one-cell context (except for unicellular organisms). If we faced a pluricellular organism, then the phenotype is the result of numerous relationships at different levels of organization, each one of them stressing different properties and pressures (figure 2). Nonetheless, genes can tell us something about a whole-organism phenotype as experiments with transgenic organisms, especially knockouts (Zan, et al. 2003), show. Its popularity and power of correlation has become so widespread in biology that it is an essential tool in any explanation effort that involves a mechanism and a model in molecular biology, genetics, neurobiology, and many others (Darden 2007). The problem with this view is that many biologists are in danger of falling, if they have not already done so, in the 'gene-for' fallacy (for an extensive and comprehensive discussion on this topic, see Kaplan and Pigliucci 2001). Godfrey-Smith (2007) also recognizes this problem and clarifies that *coding for amino acid sequences is the only semantic property that genes have*; it does not justify that genes code for

whole-organism phenotypes—they have information about them, but do not code for them. Due to this informational relationship, he likewise argues that the message is bidirectional: we can gaze something about the phenotype from the state of the genes, but we can also realize something about the genes by examining the phenotype. Nevertheless, we shall not confuse this idea of a bidirectional message with that of bidirectional transmission of information; after all, since the days of the central dogma, proteins have not yet been discovered to transmit information into genes (Maynard-Smith 2000). The notion of genes containing information about whole-organism phenotypes similarly brings into the discussion the problem of reproduction and inheritance. In turn, inheritance implies another difficulty: that an entity is both a receiver and a potential source of information—the cell that results from the reproduction of another cell is the receiver of the parent cell information; later, it is a source of information for the numerous transcription and translation processes that are going to be performed to keep the cell alive, and also for future generations of cells. Thus, information is heritable when it leads to the reconstruction of some aspects of the internal organization of one entity into another (Jablonka 2002)—inheritance becomes one of the particular features of biological information that makes it more interesting and difficult to define in terms of the metaphor.

The particularities of biological information do not end there. The instructional information as how to produce a particular effect also depends on the transcribing and translating conventions of the genetic machinery (Harms 2006). Thus, we have two more features to unfold. First, DNA also codes for the machinery responsible for its own transcription and translation. The transcription complex is composed of proteins (RNA polymerases and RNA transcription factors, among others), which are the result of the translation of some of those DNA sequences. Such translation complex is largely composed of RNAs (rRNAs and tRNAs) product of transcription, and the few proteins involved are the result of translation. Plus, that same DNA molecule codes for the proteins responsible for its replication (DNA polymerases, gyrases, helicases, single-stranded binding proteins, etc.). As a consequence, we are faced against a molecule that contains the necessary information for its own transmission and decoding, something that is not comparable to anything found in telecommunications (Shannon's main concern when proposing his mathematical theory). This brings us to the second feature: the origin and evolution of this kind of information. In addition, the question has two aspects: on the one side, the inquiry of how the translation conventions were imposed over the cellular machinery we now know; on the other, the question on the origin of the machinery responsible for replication, transcription and translation. Apparently, there should be two different answers, one for each question. Not surprisingly,

though, most well-known theories that have tried to explain the origin of biological information tackle both questions at the same time. The answer involves self-organization and natural selection ⁸ (Eigen and Maeyer 1966; Eigen 1971, Eigen and Schuster 1977; Kaufmann 1993; Schuster 2002; Popa 2004). These authors have found a way to intertwine and use them as an explanatory means to make the origin of information a more reasonable subject matter. What seems curious to me is that when one reads those discussions about the origin of biological information, one feels like reading Plato's *Cratilo* ⁹, although the technical language is more refined, equations abound, and it is not as poetic. Nonetheless, one can sense the tension between whether semantics has a convention-based origin, or if molecules themselves and their arrangements already had some kind of inherent meaning.

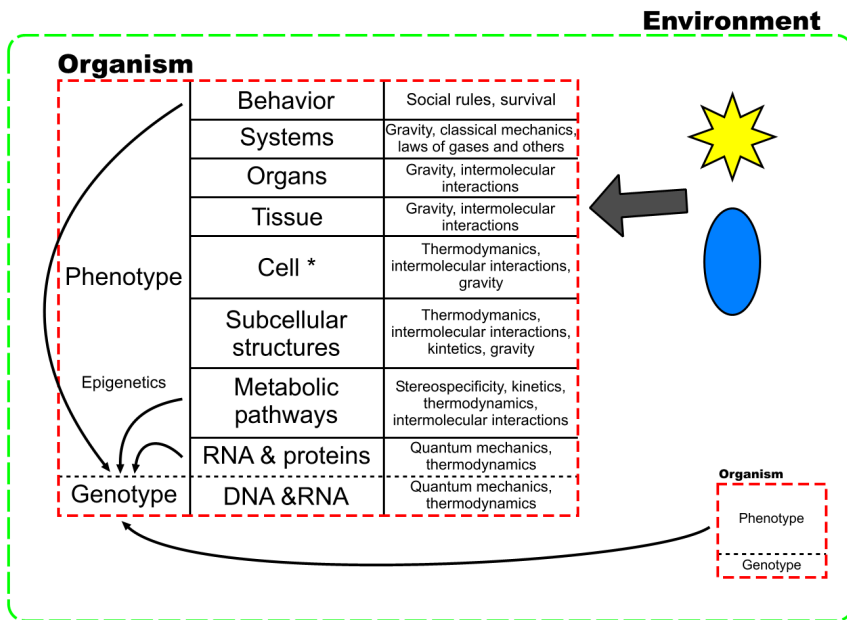


FIGURE 2
Levels of organization of the living and their interactions
(* denotes the level at which a unicellular organism arrives).

So far, we have seen some of the aspects that make semantics the prevalent view of information in biological sciences. Moreover, the last paragraph already introduced us with one of its main problems: how did the conventions or rules governing the meaning of DNA come to be? A first approach is given by the self-organization plus natural selection theories. However, what most academics have emphasized is the issue concerning intentionality that lies at the heart of this question (Millikan 1989; Maynard-Smith 2000; Jablonka 2002; Schuster 2002; Harms 2006; Godfrey-Smith 2007; Barbieri 2008b), the teleological aspect related to the purpose the sender has for the recipient within its message; an aspect that Werner Gitt (2000) called 'apobetics'¹⁰. Like 'purpose,' 'meaning' is another difficult notion in biology since it can only be defined *a posteriori* (Eigen 2000; Schuster 2002). Nonetheless, natural selection comes to the rescue. Maynard-Smith (2000) argued that the form of the signal and the response to it have evolved by selection, making genomic information meaningful as to generate an organism able to survive in the environment in which selection has acted. Therefore, intentionality, which can be understood as adaptation for Maynard Smith, becomes the result of history. A history-centered argument to approach the semantic concept of information allows to give an account of the evolutionary processes affecting the code, by which it can be comprehend how is it possible to have a false representation if the represented state is guaranteed by the occurrence of the representation. Well, it depends on the set of rules or conventions (Harms 2006), and on the fact that replication is not a 100 percent faithful. Thus, some errors appear every now and then, mutations, that alter the representation even though the set of rules has not changed. What is really interesting here is not to consider mutation, but recombination as the major source of variation, as in the case of genetic revolutions (Mayr 1954) and tinkering (Jacob 1977). That would not involve the production of new information or the removal of previous one, but the mixing of existing information, which would give us a glimpse of something else beyond the syntax and semantics of genetic information. But before diving outside of genetics, let us consider another facet of the genetic-biased view.

Like Monod (1970), one gets the feeling that Maynard Smith (2000) also considers the process of development as solely based on genetic information. He considers that genes only specify the amino acid sequence of a protein, not its three-dimensional folding, which actually depends on the laws of physics and chemistry. These laws do not have to be coded on genes because they are a given and constant part of nature. Even more, while evolution acts upon genes changing the proteins they code for, the laws of physics and chemistry remained unchanged. However, we cannot deny the importance of those laws in the correct formation of genic products—they provide the guidelines necessary for the folding and

spatial configuration of any molecule, which is certainly related to the meaning and cellular performances in which those products are involved.

4. ARE WE MISSING SOMETHING FROM BIOLOGICAL INFORMATION?

The genetic/digital-centered view of semantic information would have been enough for biology, except that the boundaries of information are much wider and, after considering that everything was already said in molecular biology, it was expected that the next frontier for information would be the nervous system (Stent 1968). Perhaps the best example of this case is F. Crick who moved from molecular biology to neuroscience in order to explore consciousness. Even if the problem of information in the nervous system and consciousness is beyond the scope of this essay, Stent's consideration is also a claim for the necessity of considering information outside a strictly genetic frame and its implications for biology. Clearly, to abandon the genetic frame implies the examination of higher levels of organization that range from the cell context to that of the whole organism, behavior and sociality. Furthermore, the understanding of information as only digital has led us to another form of reductionism: that of DNA and genes, where living organisms are nothing but DNA vessels, as seen in theories like the selfish gene (Dawkins 2006), and even practiced in recombinant DNA technology. Let us recall the case of the synthetic genome, and more scandalously, the case of the synthetic cell which many people and news headlines considered as an act of creation—what they did in fact was the *in vitro* synthesis of a chromosome and its insertion in a chromosome-less bacterial cell with all the organelles and membranes intact, delicately retained by transplantation to another cell (Gibson, et al. 2010). As Schuster (2002) stated, a claim for the creation of life would certainly have to deal with the origination of all the different parts of a cell, not just its genetic material.

When genetic information and DNA are discussed, we tend to forget the importance of all the explicit information contained in the rest of the cell (Popa 2004). Leaving aside the origin question vaguely assessed in the previous subtitle, it is necessary to consider that, although DNA possesses a great part of the information necessary to guide the development of a whole new organism, it would not be able to accomplish such gigantic task if it did not count with the molecular and cellular machinery necessary for its replication, transcription and translation. Then, the concept of biological information needs to be enlarged—the term biological information comprises not only genetic information, but also the blueprint (perhaps the only real analogy to a blueprint in the sense of an architect) information that is contained in cellular entities such as membranes or organelles, since they are also multiplied during growth and passed on to daughter cells

alongside the genetic material (Dose 1994). A similar argument can be made for the otherwise-left-aside laws of physics and chemistry. Newman, et al. (2006), argue that “an understanding of the forms assumed by multicellular organisms might not more productively be analyzed by conceptually stripping away the overlay of stabilizing and fine-tuning genetic circuitry accumulated over the last half-billion years so as to better see the originating physical and otherwise non-programmed determinants of multicellular form” (p. 289). Their ideas get even more interesting when they hypothesize that the body plans of extant organisms would be produced more or less with the same genetic toolkit, just like Jacob’s molecular tinkering, since their morphological variety would have originated due to conditional physical factors acting on viscoelastic, chemically excitable materials, and not due primarily to genetic evolution—the major role of molecular evolution over the last half billion years would have been the integration of physically inherent morphological motifs into the developmental repertoire (Newman, et al. 2006).

The history of the metaphor in developmental explanations offers another concept of information (Fox Keller 2003), and in addition Harms (2006) outlines a fifth concept. They have in common that they both deal with the fact that biological entities and information are immersed in space, a three-dimensional space in which spatial configurations are key to the understanding of their performances. The first concept is that of ‘positional information’ (Wolpert 1989). This younger brother of the much older concept of ‘organizer’ (Spemann and Mangold 1924/2001) establishes that the differentiation and pattern of gene expression of a given cell towards a certain lineage depends on its position as in a coordinate system during embryonic development. Its strongest argument was found on studies of pattern determination in *Drosophila* (Nüsslein Volhard 1991), which determined that the development of the different parts of the fruit fly in the right place depended on a series of chemical gradients of morphogens along the anterior-posterior and dorsal-ventral axes of the embryo. This research also found that most of the initial patterning determinants were mRNAs derived from the mother—maternal genes acting on the embryo. Although this notion of positional information and patterning has been extensively studied in animals¹¹, plants also display similar phenomena. Even more interesting, induced cell identity by positional information also occurs in adult plants, not only during early embryonic development. The best illustration of this case is meristems. The maintenance of the group of cells that conform the meristem of a given shoot mainly depends on a regulatory feedback loop involving two proteins: WUSCHEL (WUS) and CLAVATA3 (CLV3) (Brand, et al. 2000). WUS is a transcription factor produced by cells at the organizing center of the meristem (a region just below stem cells) that mediates the production of

an unknown signal that maintains the cells on top of the organizing center as stem cells (responsible of the continuous production of cells that further differentiate and give rise to the distinctive, and not-so-well-known, parts of a plant). In turn, *CLV3* is a protein expressed in stem cells and it indirectly suppresses *WUS* in neighboring cells, thus limiting the extension of *WUS* expression and regulating the size of stem cells region. Thus, a cell remains as a stem cell in the meristem due of its position with respect to the organizer center and to the other neighboring stem cells.

Harms's (2006) sketch of a fourth concept (in his count, but fifth in this essay) concerns Crick's phrase 'specificity of sequence' and the fact that codons are redundant: two or more codons can encode for the same amino acid although they are structurally different. Harms (2006) considers that "there is thought to be information intrinsic in a structure or form, apart from all probabilities, conventions, and causal powers" (p. 239). This time, we are faced again with the problem of space, only that it is no longer an inquiry about position, but about the structure itself. This concept would allow us to see enzymes (which Maynard Smith (2000) regarded as a non-informational source) and other higher-organization-level structures, like organs or modules, as information-bearer entities, in a way close to that of LPs—outdated discs, replaced by tapes and CDs, but greatly desired by collectors and DJs alike, that carried their melodies and songs imprinted in a series of striations in the surface of a vinyl disc that was read by a static needle while the disc revolved around its center with the needle moving from the periphery to the center thanks to the striations, while music came out of the speaker. It is a much elusive concept because of its many variants, which in turn make it hard to define. Nonetheless, it is effective at conveying the central idea of its conceptualization, although it feels more like a hunch. A classic experiment in this respect is the one performed by Skoultchi and Morowitz (1964): they used two groups of dry cysts of *Artemia spp.*; one of them was frozen at 2.2° K (or below) during six days, and the other one was a control. After the treatment, they measured the number of cysts that hatched in both groups. They found no significant difference between the two, which led them to conclude that, "Therefore the specification of a living system may be completely carried out in terms of its structure (i.e., the position of its atoms in three-dimensional coordinate space)" (Skoultchiand and Morowitz 1964, p. 163).

Root-Bernstein and Dillon (1997) have offered another way to conceptualize the structural notion through the idea of complementarity. They consider it, first and foremost, at the molecular level, so that complementarity could also be used to approach the problem of the origin of life. What makes complementarity unique is the fact that it couples all the necessary phenomena to stabilize a molecular aggregate against degradation and it also confers on all interacting parts a carrying capacity, becoming infor-

mative. Moreover, the concept of complementarity is not uniquely applicable to the molecular level, like between DNA and RNA or between an enzyme and its substrate; actually, it can also be used in higher levels of organization, such as between cells, organs and organisms, or between the organism and its surroundings, its niche. Such broad-encompassing concept allows a view of evolution as an alternating integration and divergence process, quite reminiscent of the two-step approach to evolution by natural selection.

These other forms of information call for a different definition of this matter not centered on genetic programs and DNA; one that considers any source of information based on the correspondence between source and receiver, and not focusing only on the source-capabilities of a certain entity. Eva Jablonka (2002) offers such definition:

A source—an entity or a process—can be said to have information when a receiver system reacts to this source in a special way. The reaction of the receiver to the source has to be such that the reaction can, actually or potentially, change the state of the receiver in a (usually) functional manner. Moreover, there must be a consistent relation between variations in the form of the source and the corresponding changes in the receiver¹² (p. 582).

Certainly, structural and positional information fulfills a requisite for such information concept: something could be considered an information source if it has a number of alternative states that may be detected at any particular occasion as correlated to the state of any variable that carries information concerning that source. Therefore, position could be a source of information because the fate of a given cell (the variable) during early development (or late, as we saw in plants) is correlated with the place where it might have been located; structure could be a source of information as well, as in the case of an enzyme: the state of a molecule (substrate or product) is correlated to whether the enzyme is active or not (inhibited). Godfrey-Smith (2007) criticizes such source-state/variable-state parameter for information because it would be closer to Shannon's information, regardless of semantics and, therefore, with no real value for biology. However, the critique can be surpassed if the focus on information moves from what the source of information is, to what or who is the receiver, or the 'consumer' in Millikan's (1989) terms. Millikan argued that the production of natural signs is an accidental side effect of the operation of any system. For example, if my face is red another person could think that I just blushed because I am ashamed, or that I arrived running, or maybe that I was taking the sun (for too long, I must add). Thus, the accidental signs become some sort of emergent property of the operating system for as long as they are perceived and understood by a receiver—the receiver

endows signs with specific information contents (Millikan 1989; Jablonka 2002). A receiver/consumer-centered examination of information should be focused, then, on the evolution of the system mediating the source and the receiver, i.e., the interpreting system of the receiver (Jablonka 2002). That is the essence of what Millikan (1989) called 'biosemantics'.

The broad definition of information introduced above allows for many non-genetic elements to be considered informational, as potential information sources. Recall that Maynard-Smith (2000) considered the environment as a source of noise, not of information. Anyhow, the biosemantic interpretation would present us an environment ¹³ full of information-bearer entities or factors, as the many studies on epigenetics have shown ¹⁴ (Jablonka 2002). Environmentally-elicited responses on pluricellular organisms, such as animals and plants, have been studied for centuries; even unicellular organisms show important cues, as in bacteria: from the elder *lac operon*, which depended upon lactose, an abiotic external factor, to the almost pubescent 'quorum sensing' (Waters and Bassler 2005), which depends on the presence of other bacteria giving rise to a density-dependent population effect. Moreover, by mixing together the definition of information given by Jablonka and her considerations on what makes information heritable (reviewed in the last section), she arrives at the conclusion that there are other inheritance systems apart from the genetic one: epigenetic (cellular heredity, not only within an organism, but from one organismal generation to the next); behavioral (behavioral preferences and patterns), and symbolic (transmission of information through symbols and its role on cultural and social evolution, specifically human). (For an extensive and careful examination of these inheritance systems, I suggest the reading of Jablonka and Lamb 2005 ¹⁵.) There is a question about information, one that I am not going to assess here but seems important to bring it up, that arises from the very first definition in Shannon's terms and acquires even more validity in the light of this latter definition: Do we need a new physical ontology in order to understand information as another fundamental quantity of the universe, alongside matter, energy and time?, or is it an epistemological problem?

It is common to find conflicting views in any group and, instead of being an exception, the scientific community could be the flag of this claim. As the reader may have noticed so far, this essay has nurtured on this feature, on the many discrepancies found among scientists and philosophers, and it will continue to do so. Millikan (1989) and Jablonka's (2002) biosemantics have found an adversary in a different communicative view of life, biosemiotics. They have in common that they are not limited to a digital form of information, but the name itself underlies one difference: while biosemantics focuses on meaning, biosemiotics focuses on the uses and functions of signs and codes (Barbieri 2008a), indifferent of meaning. Thus,

by taking meaning out of the inquiry the problem of information moves its focus back to the source and to the coding capabilities of any living system—to the codemaker instead of the interpreter (Barbieri 2008b). In spite of how exciting this new biological perspective may look and the frontiers it pushes, prevent me, for the sake of sanity—mine as well as the reader's—from enlarging even more this never-ending discussion, and allow me to recapitulate and conclude.

5. CONCLUSIONS

The metaphor of information in biological sciences was born mainly as an issue of genetics and development, solidly sustained in molecular grounds. As a metaphor, its only defined property is to fill in the gaps observed within theories (Fracchia and Lewontin 2005), in this case, the gap of cell differentiation during the 1950s and 1960s. Later, new findings, discoveries, theories, and models changed the course of the metaphor in ways that does not fit so well with the source of such metaphor. Nevertheless, it is a metaphor that has proven invaluable for biological sciences (Godfrey-Smith 2007). Despite its many faces and concepts, we can be sure that the issue on biological information has to take into account the fact that, regardless what concept we are using, it will always have to deal with its transmission—it is the only common feature shared by all the concepts explored here. As a consequence, any effort towards a wide conceptualization of information would have to start by exploring that which is transmissible within a living system (this consideration would imply different consequences for unicellular or multicellular organisms, especially since Weismann's germ-some differentiation is so widely accepted for the latter, and not so much for the former, especially when discussing bacteria). In addition, the many turns it has experienced make the blurry concept of biological information more suitable for the land of hi-fi and HD that biological sciences explore: A land in which any communication process, including heredity, can be classified as 'high-fidelity,' yet leaving room to imperfection and evolution; and a land in which 'high definition,' in a three-dimensional context, can be found throughout all its organization levels, from the digital and molecular realm of nucleic acids and proteins, to the more familiar realm of organisms and ecosystems.

ACKNOWLEDGEMENTS

I would like to thank to professors Gloria Machado and Sergio Orozco, my tutors in this dissertation, for their support throughout this process; to the Comité para el Desarrollo de la Investigación (CODI), from Universidad de Antioquia, for the financial aid, and to professor Eugenio Andrade for his invaluable support, comments, and discussions.

NOTES

- 1 Nonetheless, the relationship between information theory and computer science is inherent to them since most of computer science is written in the language of bits.
- 2 The notion of endpoint is directly related to the concept of teleonomy, the apparent goal-seeking behavior observed in living systems. That concept is crucial in Mayr's philosophy of biology and his advocacy for the autonomy of biology as a science. However, this concept will not be assessed in this essay.
- 3 We could currently point out a difference that would have saved Mayr's conceptualization: the genome also encompasses both the coding as well as the non-coding DNA sequences. The genetic program would be located only in the coding DNA sequences. However, he never refers to it, not even in his last publication.
- 4 Nevertheless, the concept itself of genetic program is never thoroughly defined; it remains more like the metaphor it is than as an actual concept of biological sciences. See Fox-Keller (2003) for a wider discussion.
- 5 Shannon (1948) defines these five parts like this:
Information source: produces a message or sequence of messages to be communicated to the receiving terminal.
Transmitter: operates on the message in some way to produce a signal suitable for transmission over the channel.
Channel: the medium used to transmit the signal from transmitter to receiver.
Receiver: performs the inverse operation of that done by the transmitter, reconstructing the message from the signal.
Destination: the person (or thing) for whom the message is intended.
- 6 This analogy is a highly controversial topic, having many defendants and many detractors (Harms 2006). Because it is not within the scope of this essay, I will summarize the concept without validating or questioning it.
- 7 Notice that Maynard Smith deals only with the problem of translation. Thus, the only products can be proteins. The RNAs mentioned before are the product of transcription only and other post-transcriptional modifications not related to translation.
- 8 The exact way how this two concepts interplay to give rise to biological information is not part of this discussion as I am dealing here with the concept and its uses in biological sciences given that it is a borrowed metaphor to explain some observations of biology, not with information itself and its history in the origins of life. If the reader is interested in such matters, I recommend her to check the references listed there.

- 9 Take for example this quote: “Well, Socrates, I’ve often talked with Cratylus—and with lots of other people, for that matter—and no one is able to persuade me that the correctness of names is determined by anything besides convention and agreement. I believe that any name you give a thing is its correct name. If you change its name and give it another, the new one is as correct as the old. For example, when we give names to our domestic slaves, the new ones are as correct as the old. No name belongs to a particular thing by nature, but only because of the rules and usage of those who establish the usage and call it by that name. However, if I’m wrong about this, I’m ready to listen not just to Cratylus but to anyone, and to learn from him too” (Plato, *Cratylus*, 384c-e).
- 10 Gitt is a creationist who has written several papers concerning information theory. His argument has focused on the use of information, and especially apobetics, as a supporting element of creationism.
- 11 See the case of homeobox genes in Duboule 2000.
- 12 After the definition, Jablonka continues, “What I mean by ‘form’ is the organization of the features and/or the actions of the source, and specifically those aspects of organization with which the receiver reacts. The source eliciting the special type of reaction in the receiver will be referred to as ‘input’ or ‘information cue,’ and the processes in the receiver that result in a regular and functional response will be called ‘interpretation.’ The term ‘signal’ will be reserved only for evolved informational inputs, that is, evolved inputs produced by an evolved or otherwise designed source. ‘Functional’ is here used to mean the consistent causal role that a part plays within an encompassing man- designed or natural-selection-designed system, a role that usually contributes to the goal-oriented behavior of this system.” Note that she also refers to function in this definition. Although function is a central concept of biology and teleological interpretations of the living, a discussion about it would take a whole new essay. I shall turn my back on it, then, and leave it aside.
- 13 Here, I take the definition of environment as in Lewontin (2004): Aspects of the external world that become relevant when functionally associated to an organism’s activities.
- 14 For a reference book on epigenetics and the environment see Gilbert and Epel 2009.
- 15 Because this essay moves within a Darwinian frame, I consider important to make this annotation. Some colleagues and I have had some discussions regarding Jablonka and Lamb’s book. The main disagreement is whether they are arguing against a Darwinian model of evolution to favor what some have called Neo-Lamarckism, or expanding the scope of Darwinism. However, after reading Jablonka (2002) one can be sure that the main argument of her later collaborative work is not to defy natural selection and Darwinism, but to argue against the narrow view imposed on information and the importance of their wide conception in evolution and the new doors it opens for understanding its processes, including natural selection and the origin of the variations upon which selection acts. The alternative systems described by Jablonka and Lamb could be encompassed within the group of variation sources alongside mutation and recombination, upon which natural selection could act—it broadens the traditional conception without denying the importance of natural selection in the process of evolution.

REFERENCES

- Barbieri M. (2008a), "Biosemiotics: A new understanding of life," *Naturwissenschaften* 95(7): 577-599.
- (2008b), "Is the cell a semiotic system?," in Barbieri M (Ed). *Introduction to Biosemiotics: the New Biological Synthesis*. Dordrecht: Springer, pp. 179-208
- Baierlein R. (2003), *Thermal Physics*. Cambridge: Cambridge University Press.
- Berg P.; Singer M. F. (1995), "The recombinant DNA controversy: twenty years later," *PNAS* 92(20): 9011-9013.
- Berg P., et al. (1975), "Summary statement of the Asilomar Conference on recombinant DNA molecules," *PNAS* 72(6): 1981-1984.
- Brand U., et al. (2000), "Dependence of stem cell fate in *Arabidopsis* on a feedback loop regulated by CLV3 activity," *Science* 289(5479): 617-619.
- Crick F. H. (1970), "Central Dogma of molecular biology," *Nature* 227(5258): 561-563.
- Darden L. (2007), "Mechanisms and models," in Hull D, Ruse M (eds.), *The Cambridge Companion to the Philosophy of Biology*. Cambridge: Cambridge University Press, pp. 139-159.
- Dawkins R. (2006), *The Selfish Gene*. New York: Oxford University Press.
- Dose K. (1994), "On the origin of biological information," *Journal of Biological Physics* 20(1): 181-192.
- Duboule D. (2000), "Developmental genetics: A Hox by any other name," *Nature* 403(6770): 607-610.
- Eigen M. (1971), "Self-organization of matter and the evolution of biological macromolecules," *Naturwissenschaften* 58(10): 465-523.
- (2000), "Natural selection: a phase transition?" *Biophysical Chemistry* 85(2-3): 101-123.
- Eigen M.; Maeyer L. (1966), "Chemical means of information storage and read-out in biological systems," *Naturwissenschaften* 53(3): 50-57.
- Eigen M.; Schuster P. (1977), "A principle of natural self-organization," *Naturwissenschaften* 64(11): 541-565.
- Fox Keller E. (2003), *Making Sense of Life: Explaining Biological Development with Models, Metaphors and Machines*. New York: Oxford University Press.
- Fracchia J.; Lewontin R. C. (2005), "The price of metaphor," *History and Theory* 44: 14-29.
- Gibson D., et al. (2010), "Creation of a bacterial cell controlled by a chemically synthesized genome," *Science* 183(127): 810-816.
- Gilbert S. F.; Epel D. (2009), *Ecological Developmental Biology: Integrating Epigenetics, Medicine, and Evolution*. Sunderland: Sinauer Associates.
- Gitt W. (2000), *In the Beginning was Information*. Berlin: CLV.
- Godfrey-Smith P. (2007), "Information in biology," in: Hull D, Ruse M (Eds). *The Cambridge Companion to the Philosophy of Biology*. Cambridge: Cambridge University Press, pp. 103-119.
- Harms W. F. (2004), *Information and Meaning in Evolutionary Processes*. Cambridge: Cambridge University Press.
- (2006), "What is information? Three concepts," *Biological Theory* 1(3): 230-242.
- Jablonka E. (2002), "Information: Its interpretation, its inheritance, and its sharing," *Philosophy of Science* 69(4): 578-605.
- Jablonka E.; Lamb M. (2005), *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral and Symbolic Variation in the History Of Life*. Cambridge: MIT Press.
- Jacob F. (1970), *La logique du vivant*. Paris: Gallimard.
- (1977), "Evolution and tinkering," *Science* 196(4295): 1161-1166.

- Jacob F.; Monod J. (1961), "Genetic regulatory mechanisms in the synthesis of proteins," *Journal of Molecular Biology* 3: 318-356.
- Kaplan J.; Pigliucci M. (2001), "Genes 'for' phenotypes: A modern history view," *Biology and Philosophy* 16(2): 189-213.
- Kauffman S. (1993), *The Origins of Order*. New York: Oxford University Press.
- Kimura M. (1968), "Evolutionary rate at the molecular level," *Nature* 217(5129): 624-626.
- Lewontin R. C. (2004), "Building a science of population biology," in: Singh R. S.; Uyenoyama M. K. (Eds). *The Evolution of Population Biology*. Cambridge: Cambridge University Press, pp. 7-20.
- Lilley D. (2003), "The origins of RNA catalysis in ribozymes," *Trends in Biochemical Sciences* 28(9): 495-501.
- Maynard Smith J. (2000), "The concept of information in biology," *Philosophy of Science* 67(2): 177-194.
- Mayr E. (1954), "Evolution and change in genetic environment," in: Huxley J, Hardy A. C., Ford E. B. (Eds). *Evolution as a Process*. New York: Allen & Unwin, pp. 157-180.
- (1961), "Cause and effect in biology," *Science* 134: 1501-1506.
- (1974), "Teleological and teleonomic: A new analysis," *Boston Studies in the Philosophy of Science* 14: 91-117.
- (1982), *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*. Cambridge: The Belknap Press of Harvard University Press.
- (1992), "The idea of teleology," *Journal of the History of Ideas* 53(1): 117-135.
- (2004), *What Makes Biology Unique. Considerations on the Autonomy of a Scientific Discipline*. Cambridge: Cambridge University Press.
- Millikan R. G. (1989), "Biosemantics," *The Journal of Philosophy* 86(6): 281-297.
- Monod J. (1970), *Le Hassard et la Nécessité. Essai sur la philosophie naturelle de la biologie modern*. Paris: Éditions du Seui
- Morange M. (2005), "What history tell us: I. The operon model and its legacy," *Journal of Biosciences* 30(3): 313-316.
- Newman S. A., et al. (2006), "Before programs: the physical origination of multicellular forms," *International Journal of Developmental Biology* 50(2-3): 289-299.
- Noller H. F. (2005), "RNA structure: reading the ribosome," *Science* 309 (5740): 1508-1514.
- Nüsslein-Volhard C. (1991), "Determination of embryonic axes of *Drosophila*," *Development Supplement* 1: 1-10.
- Popa R. (2004) *Between Necessity and Probability: Searching for the Definition and Origin of Life*. Dordrecht: Springer-Verlag.
- Root-Bernstein R.S., Dillon P. F. (1997), "Molecular complementarity I: the complementarity theory of the origin and evolution of life," *Journal of Theoretical Biology* 188(4): 447-479.
- Schuster P. (2002), "The origin and processing of biological information," in: Baltimore D., et al. (eds). *Frontiers of Life, Volume 1*. New York: Academic Press, pp. 329-346.
- Shannon C. (1948), "A mathematical theory of communication," *The Bell System Technical Journal* 27: 379-423.
- Skoultchi A.I., Morowitz H. J. (1964), "Information storage and survival of biological systems at temperatures near absolute zero," *The Yale Journal of Biology and Medicine* 37: 158-163.

- Smith H. O., et al. (2003), "Generating a synthetic genome by whole genome assembly: phiX174 bacteriophage from synthetic oligonucleotides," *PNAS* 100(26): 15440-15445.
- Spemann H.; Mangold H. (1924/2001), "Induction of embryonic primordial by implantation of organizers from a different species," *The International Journal of Developmental Biology* 45(1): 13-38.
- Stent G. (1968), "That was the molecular biology that was," *Science* 160(826): 390-395.
- Trevors J.; Abel D. (2004), "Chance and necessity do not explain the origin of life," *Cell Biology International* 28(11): 729-739.
- Waters C. M.; Bassler B. I. (2005), "Quorum sensing: cell-to-cell communication in bacteria," *Annual Reviews of Cell and Developmental Biology* 21: 319-346.
- Wolpert L. (1989), "Positional information revisited," *Development* 107 (supplement): 3-12.
- Zamore P. D. (2006), "RNA interference: big applause for silencing in Stockholm," *Cell* 127(6): 1083-1086.
- Zan Y., et al. (2003), "Production of knockout rats using ENU mutagenesis and a yeast-based screening assay," *Nature Biotechnology* 21(6): 645-651.