BioSystems 181 (2019) 11–19

Contents lists available at ScienceDirect

BioSystems

journal homepage: www.elsevier.com/locate/biosystems

A general model on the origin of biological codes

Marcello Barbieri

Dipartimento di Morfologia ed Embriologia, Via Fossato di Mortara 64a, 44121 Ferrara, Italy

ARTICLE INFO

Keywords:
Organic codes
Neural codes
Cultural codes
Arbitrariness
Code ambiguity
Code optimization
Code conservation
Major transitions

ABSTRACT

For a long time it has been assumed that the rules of the genetic code were determined by chemistry – either by stereochemical affinities or by metabolic reactions – but the experimental evidence has revealed a totally different reality; it has been shown that any codon can be associated to any amino acid, and this means that there is no deterministic link between them. The genetic code, in other words, is based on *arbitrary*, or *conventional*, rules and this raises a formidable problem: how can arbitrary rules exist in Nature? We know that such rules exist in culture, but there is an abyssal difference between biology and culture, because the cultural codes are short-lived, whereas the biological codes are the most conserved entities in evolution. Biological codes, in other words are fundamentally different from cultural codes and we do need a model that makes us understand how they came into being.

In this paper it is shown that the origin of biological codes takes place in five phases (beginning, evolution, optimization, major transition and conservation) and this suggests a general model for their development. According to this model, a biological code evolves in a system as a means of solving a local problem, but then it becomes the tool of a much larger change in macroevolution. This is the great potential of the biological codes: their ability to bring into existence absolute novelties that change the whole course of the history of life. Different major transitions were based on different codes, but we can also recognize some common features in all of them. This indicates that coding is a universal mechanism that Nature has employed many times in the course of evolution to solve a wide variety of different problems.

1. Introduction

In Life Itself (1981) Francis Crick wrote that "… the genetic code is as important for biology as Mendeleev's Periodic Table of the Elements is for chemistry, but there is an important difference. The Periodic Table would be the same everywhere in the universe. The genetic code appears rather arbitrary, or at least partly so … If this appearance of arbitrariness in the genetic code is sustained, we can only conclude that all life on earth arose from one very primitive population" (Crick, 1981, p 46–47).

A few years later that "*appearance of arbitrariness*" became a certainty because it was shown that any codon can be associated with any amino acid (Schimmel, 1987; Schimmel et al., 1993; Budisa, 2004; Hartman et al., 2007) thus proving that there are no deterministic links between them.

The origin of the genetic code is a historical fact, like the origin of the solar system, the origin of the stars and the origin of the universe, but in this case we must take into account the important difference outlined by Crick, the fact that we are in the presence of *arbitrary rules*, and that raises a problem that science has never met before: how can arbitrary rules appear in Nature?

If the genetic code were an isolated case we could put this problem aside, but in the past twenty years it has been shown that many other organic codes exist in living systems (Barbieri, 2003), so how do we account for their existence in Nature? It is true that each code is unique, but the fact that many of them appeared on Earth tells us that Nature has resorted to the *mechanism of coding* time and time again in the history of life, and this is something that we need to understand.

Here it is shown that a general model on the mechanism of coding can be obtained by studying the most general characteristics of the genetic code; more precisely, it is shown that the history of the genetic code can be divided into five phases that do not depend on the specific characteristics of that code, and this gives us a model that can be applied to other codes. The five phases can be described in this way:

1.1. Beginning

The first rules of a code appear in a system as a means of performing a particular function, and their arbitrariness implies that the first version of a code is necessarily ambiguous.

E-mail address: brr@unife.it.

https://doi.org/10.1016/j.biosystems.2019.04.010
Received 22 March 2019; Received in revised form 16 April 2019; Accepted 16 April 2019
Available online 17 April 2019
0303-2647/

1.2. Evolution

If the products of an ambiguous code can still play a useful role in a system, evolution tends to improve them by steadily reducing the ambiguity of the code.

1.3. Optimization

During the evolution of a code, there is a phase in which the system can submit its rules to a process of optimization.

1.4. Major transition

When the ambiguity of a code is completely eliminated, the system starts producing components that have biological specificity and set in motion a major transition in evolution.

1.5. Conservation

After a major transition the system continues to evolve and change but the rules of the code are transmitted virtually intact to the descendants and are highly conserved in evolution.

The value of this model is that the above five phases can be observed in other cases as well, and in particular in the history of the neural code, the set of rules that the brain employs to transform the signals from the sense organs into instincts and feelings.

The present paper has the goal to illustrate a general model on the origin of biological codes and to this purpose it is divided into three parts: the first is dedicated to the genetic code, the second to the neural code and the third to the relationships that exist between codes and macroevolution.

2. Part 1: the genetic code

2.1. Genetic code phase 1 – beginning

A code is a set of rules that a system employs to perform a particular function, and its history necessarily takes place within that system. The genetic code is the set of rules that are used in protein synthesis and its history necessarily began when the first apparatus of protein synthesis appeared on the primitive Earth. Today protein synthesis is based on three types of RNAs: the messenger-RNAs provide a sequence of codons, the transfer-RNAs associate codons to amino acids and the ribosomal-RNAs join amino acids together into proteins. The characteristics of the modern RNAs, furthermore, allow us to reconstruct at least some of the characteristics of their ancestral precursors.

2.1.1. Ancestral ribosomal-RNAs

The ribosomal RNAs are among the most conserved molecules in evolution (Woese, 1987, 2000) and contain regions that have the ability to form peptide bonds (Nitta et al., 1998). This suggests that the ribosomal-RNAs appeared very early on the primitive Earth and some of them could stick amino acids together in no prefixed order thus producing random proteins.

2.1.2. Ancestral transfer-RNAs

The modern transfer-RNAs are small molecules that deliver amino acids to the ribosomal-RNAs and have a basic cloverleaf structure that has been highly conserved in evolution, a clear sign that they descended from a common ancestor; it must be underlined, however, that their contribution to protein synthesis required the presence of a third type of RNAs.

2.1.3. Ancestral anchoring-RNAs

When amino acids are delivered by transfer-RNAs to the site of protein synthesis, it is necessary that they are kept at a close distance for a long enough time to allow the formation of a peptide bond (Wolf and Koonin, 2007; Fox, 2010). This means that the transfer-RNAs must have temporary anchoring sites, and in primitive systems these were provided by *anchoring-RNAs*, the ancestors of the modern *messenger-RNAs*.

The combination of ribosomal-RNAs, transfer-RNAs and anchoring-RNAs gave origin to an apparatus of protein synthesis where the transfer-RNAs were automatically creating a *mapping* between codons and amino acids, and any such mapping is, by definition, a *genetic code*. This is how the genetic code came into being: that code appeared on Earth when ancestral transfer-RNAs and anchoring-RNAs joined the ancestral ribosomal-RNAs and became an integral part of the ancestral apparatus of protein synthesis.

The key point is that the first genetic code was necessarily *ambiguous* because nothing could prevent a codon from coding two or more amino acids (Fitch and Upper, 1987; Osawa, 1995). This means that a sequence of codons was translated some time into a protein and some other time into a different protein, and the apparatus was inevitably producing *statistical* proteins.

Statistical proteins cannot act as specific enzymes but can still perform useful functions. In the ancestral systems it was the RNAs that were acting as specific enzymes but they could barely work on their own. They needed peptides and polypeptides to maintain stable conformations in space and this *ancillary* function did not require proteins with specific sequences.

Statistical proteins, furthermore, could arise from various combinations of acidic, basic, hydrophilic and hydrophobic amino acids, and this gave them the ability to form helices and pleated sheet, to separate hydrophilic from hydrophobic regions, to assemble themselves into supramolecular structures, to create microenvironments and above all to maintain the system in a state of continuous metabolic activity.

In primitive systems that were producing statistical proteins, in short, a fraction of these proteins were likely to perform useful functions. But could those systems evolve?

2.2. Genetic code phase 2 – evolution

Protein synthesis requires three types of RNAs, but the molecular machines that actually make proteins, the *ribosomes*, are made of ribosomal-RNAs and more than 50 ribosomal proteins, each present in one or a few copies (Nomura et al., 1974). The ancestral systems, on the other hand, could only produce *statistical* proteins, and one may wonder whether their ribosomes could work with such proteins. Surprisingly, this is a question to which we can give a positive answer.

A first reason is the fact that the ribosomal proteins change from species to species, which means that the same function can be performed by different proteins. In bacteria, furthermore, the ribosomal proteins are fewer and smaller than in eukaryotes, which means that their number and their molecular weights are not critical factors. The ribosomal proteins, in other words, can have different weights, different numbers and different compositions and yet they can all form fully operational ribosomes. The ancestral ribosomes, in short, did not require specific proteins and could well have worked with statistical proteins.

At this point we have the problem of understanding what types of statistical proteins made up the ancestral ribosomes and here an important clue has come from self-assembly experiments. Ribosomes are formed by self-assembly from their components, and it has been possible to discover the contribution of individual ribosomal proteins by studying what happens when ribosomes are reassembled without anyone of them in turn. These experiments have shown that the ribosomal proteins fall into three major groups: some are necessary for function, others are required for self-assembly, and those of the third group have a stimulating effect but are fundamentally disposable (Kurland, 1970; Fox, 2010).

The evolution of the ribosomal proteins was therefore a process that gave origin first to three great families and then to an increasing number of protein subfamilies. The ancestral ribosomes could work with statistical proteins but there was one condition that had to be met: they could survive in evolution only if the same families and subfamilies of statistical proteins could reappear in every new generation.

Let us examine, for example, the self-assembly of the ancestral ribosomes. This process was repeated in every generation, but when the genetic code was ambiguous the proteins involved in self-assembly were affected by that ambiguity and this had an outstanding consequence: the same families of statistical proteins could reappear in the descendants only if the ambiguity of the genetic code was not cancelling out their family differences.

The ambiguity of the genetic code, in other words, was the limiting factor that determined how many families of statistical ribosomal proteins could reappear in the descendants. Which means that only by lowering the ambiguity of the code it was possible to promote the evolution of the ribosomal proteins. An evolution that increased their number, diversified their functions and favoured their self-assembly in every new generation (Barbieri, 2015b).

The purpose of the ancestral systems was not the synthesis of specific proteins because they could not evaluate the *future* benefits of such proteins. What they could evaluate, however, were the *immediate* benefits of ribosomal machines that were increasingly efficient in producing their statistical proteins, and it was for this reason that evolution systematically decreased the ambiguity of the ancestral genetic code.

It will be noticed that the evolution of the ribosomal proteins was not about this or that protein or this or that protein function. It was involving all ribosomal proteins at the same time. It was not about *individual* features but about *collective* relationships. It was the evolution of a systems as a whole, an evolution that went on until the ambiguity of the genetic code was completely removed and *biological specificity* in protein synthesis came into existence.

2.3. Genetic code phase 3 – optimization

The modern genetic code is a mapping between 64 codons carried by transfer-RNAs and 20 amino acids carried by 20 aminoacyl-tRNA-synthetases, each of which attaches one amino acid to one or more tRNAs. The synthetases are specific proteins that can be produced only by an apparatus that already has a genetic code, and this gives us a classic *chicken-and-egg* paradox: how could the genetic code come into existence if its rules are implemented by proteins that can be made only when the code already exists?

A possible solution is that the *modern* apparatus of protein synthesis was preceded by an *ancient* apparatus where the amino acids were attached to the transfer-RNAs not by proteins but by RNAs (Maizels and Weiner, 1987). The *modern genetic code*, in other words, was preceded by an *ancient genetic code* based on *RNA-synthetases* that were later replaced by *protein-synthetases*.

Such a replacement, on the other hand, was bound to have major biological consequences. Proteins can mimic RNAs but only up to a point, and replacing the RNA synthetases with protein synthetases could well have modified the rules of the ancient genetic code. But did that actually happen?

The evidence in support of the idea that the ancient code has been repeatedly modified has come from computer simulation studies which suggest that the modern code performs better than most of its potential alternatives (Haig and Hurst, 1991; Freeland and Hurst, 1998; Bollenbach et al., 2007). There is however some disagreement about the characteristics that have been optimized.

According to Carl Woese, the genetic code has been optimized for minimizing the impact of translation errors (Woese, 1965; Woese et al., 1966), whereas Gilis et al. (2001) have proposed that the modern code is optimal in respect to the stabilization of protein structure. Itzkovitz and Alon (2007) have argued that the modern code is nearly optimal for the acquisition of additional information into genetic sequences, whereas Drummond and Wilke (2008) have suggested that the modern code is ideally suited to favor the process of protein folding. On the whole, the computer simulations support the idea that the genetic code went through processes of optimization, and we cannot exclude the possibility that it was simultaneously optimized for a variety of different properties.

It must be underlined that some authors have warned against reaching overoptimistic conclusions on this issue. Novozhilov et al. (2007) have pointed out that there are 10^{84} possible codes and many of them are more robust than the modern code. Their computer simulations have revealed that the genetic code did go through processes of optimization but apparently it went only half way up the optimality ladder.

Whatever did happen in the optimization phase, it seems that the introduction of protein synthetases in protein synthesis not only replaced the ancient genetic code with the modern one, but provided the conditions for optimizing its performance, a process that gave origin to a genetic code of extreme virtuosity and the accuracy of protein synthesis became so high as to be virtually error-free.

2.4. Genetic code phase 4 – major transition

When the ambiguity of the genetic code was completely eliminated, the ancestral systems acquired the ability of producing specific proteins. At that point they could have employed these proteins for the same functions as the previous statistical proteins, but in reality they did something completely different. They started using specific proteins for entirely new functions or for functions that were previously performed by the RNAs, and in this way initiated a revolution that eventually transformed the ancient RNA world into the modern protein world. Once in existence, in other words, the genetic code set in motion a massive change in macroevolution, a true major transition in the history of life.

We have already seen a first example of the protein revolution in the change that occurred in protein synthesis when the protein-synthetases replaced the previous synthetases, a change that transformed the ancient genetic code into the modern code.

A second example comes from the *ATP synthase*, a molecular machine that represents one of the few universals of biology, like DNA, ribosomes and the genetic code. Virtually all cells obtain energy from proton currents generated by *ATP synthase*, a machine that contains some two dozen proteins organized into four functional elements (headpiece, rotor, shaft and stator) (Harold, 2014). These *specific* proteins could be synthesized only after the appearance of the modern genetic code, and yet even the previous ancestral systems needed energy and could only get it by similar energy-transduction means. This is why it has been proposed that the modern ATP synthases took the place of previous RNA machines that were operating in the ancestral RNA world (Mulkidjanian et al., 2007).

Another outstanding example of the protein revolution comes from the machinery of DNA replication. One can hardly imagine the ancestral systems without replicating DNA molecules, and yet most of the enzymes involved in DNA replication are radically different in bacteria and archaea, which means that the modern mechanisms of DNA replication have been re-invented at least twice by the descendants of the common ancestor (Woese, 2002; Koonin and Martin, 2005).

The transition from the ancient RNA world to the modern protein world started with the appearance of the modern genetic code in the last universal common ancestor, and went on until the origin of the first modern cells, when the transition to protein metabolism was virtually complete and life as we know it came into being. The characteristics of the ancestral cells have been reconstructed by molecular phylogenetic trees and are traditionally represented by the three primary kingdoms that Carl Woese called Archaea, Bacteria and Eukarya (Woese, 1987, 2000). It must be underlined, however, that the origin of the first cells was not exclusively due to the protein revolution. Another major contribution came from the development of different types of cell membranes in the descendants of the common ancestor (Lombard and Moreira, 2011).

The cell membrane has a special role in life because it is never constructed *de novo*. Membranes always grow from pre-existing membranes, and this has lead to the concept of *membrane heredity*, the idea that membranes are passed down from one generation to the next in an uninterrupted chain of descent (Blobel, 1980; Sapp, 1987; Cavalier-Smith, 2000; Harold, 2005). Chromosomes are reproduced from pre-existing chromosomes and membranes from pre-existing membranes but they carry two very different types of instructions. Chromosomes transmit genetic information, whereas membranes transmit architectural order.

We realize in this way that the major transition of the protein revolution went on in parallel with the membrane revolution, and both took place in the descendants of the common ancestor after the origin of the modern genetic code.

2.5. Genetic code phase 5 – conservation

The genetic code is the set of rules that the apparatus of protein synthesis employs to make proteins, and yet there is a profound difference between the evolution of the genetic code and the evolution of the apparatus of protein synthesis. The genetic code has been highly conserved since its origin almost 4 billion years ago, whereas the apparatus of protein synthesis has continued to evolve and to change.

In prokaryotes, for example, the ribosomes have molecular weights of about 2 million, whereas in eukaryotes their weights are more than 4 millions. In bacteria the number of ribosomal proteins is 57, in archaea 68 and in eukaryotes 78 (Lecompte et al., 2002). In eukaryotes, the ribosomal RNAs are much heavier than in prokaryotes, and the biogenesis of the ribosomes is extremely more complex. In fact, virtually everything has been modified in the apparatus of protein synthesis in the course of evolution and the sole outstanding exception is the rules the genetic code. They are the sole entities that have been conserved for billions of years while everything else has been changed.

This extraordinary process of conservation is usually accounted for by saying that the genetic code is a set of *constraints* (Pattee, 2001; Gould, 2002) and that physical constraints cannot be changed, an idea that appears to explain why the genetic code has been *frozen* since the origin of life.

The statement that the genetic code is a set of constraints is *formally* correct because its rules impose severe limitations on a virtually unlimited number of possibilities. It must be underlined, however, that they are not *physical* constraints. A big piece of rock on a road is a physical constraint, but a traffic light is not; a traffic light is a code, a totally different type of constraints.

The genetic code is made of *biologically generated* rules that in no way can be assimilated to physical constraints because the genes of the molecules that implement the genetic code are constantly subject, like all other genes, to mutation and neutral drift. They are in a continuous state of flux and the fact that they have been highly conserved in evolution means that there is a biological mechanism that *actively and continuously* restores their original structure.

The conservation of the genetic code, in other words, is not the result of physical constraints but of active biological mechanisms that are continuously at work. It is tempting to say that these mechanisms consists in the standard processes of gene replication and gene repair, but these processes work in the same way on all genes and do not explain why some genes are more conserved than others.

From a theoretical point of view, the conservation of the genetic code is studied in the framework of *autopoiesis*, the idea that living systems have the ability to produce their own components and to conserve them in time (Maturana and Varela, 1980). In reality, most components of the living systems can be changed in the long run, and only a few must be absolutely conserved. More precisely, it is the coding rules that are highly conserved in evolution, so what we have is not autopoiesis but *codepoiesis* (Barbieri, 2012). The conservation of the genetic code, in conclusion, is an experimental reality but it is also a major theoretical problem that remains to be solved.

3. Part 2: the neural code

3.1. Introduction to the neural code

There is a large consensus today that mind is a natural phenomenon and that mental events are produced by brain events. More precisely, it is widely accepted that mind is made of higher-level brain processes, such as feelings and instincts, that are produced by lower-level brain processes such as neuron firings and synaptic connections (Searle, 2002). We need therefore to understand *how* does the brain produce the mind and today the major scientific theories that have been proposed on this issue can be divided into four groups.

(1) The *computational theory* is the idea that lower-level brain processes are transformed into feelings and instincts by neural processes that are equivalent to computations (Fodor, 1975, 1983; Johnson-Laird, 1983).
(2) The *connectionist theory* maintains that the brain is solving problems by means of neural networks that operate with explorative strategies, and feelings and instincts arise as side-effects of these networks (Hopfield, 1982; Rumelhart and McClelland, 1986; Edelman, 1989; Holland, 1992; Churchland and Sejnowski, 1993; Crick, 1994).
(3) The *emergence theory* states that higher-level brain properties emerge from lower-level neurological phenomena, and mind is distinct from brain because any emergence is accompanied by the manifestation of new properties (Morgan Lloyd, 1923; Searle, 1980, 2002).
(4) The *code theory* is the idea that there has been a neural code at the origin of mind as there has been a genetic code at the origin of life (Barbieri, 2003, 2011, 2015a,b).

All four theories have been inspired by observations and have the goal of describing some real aspects of brain activity; neural computations and neural networks, for example, are eminently suited to illustrate the automatic operations of the *unconscious* brain, and we should not forget that these operations account for the vast majority of brain activities. It must be underlined, however, that there is a substantial difference between the code theory and the other three.

According to the code theory, feelings and instincts are not the side-effects of neural networks (as in *connectionism*), they do not come into existence by *emergence*, and they are not the result of *computations*. They are *manufactured* from lower-level brain processes according to the rules of the neural code just as proteins are manufactured from amino acids according to the rules of the genetic code. In the framework of the code theory, in other words, feelings and instincts are *manufactured brain artifacts*, whereas according to the other theories they are *spontaneous brain products*.

The code theory, in conclusion, is a new approach to mind and, as we will see, there is a substantial body of evidence in its favor, so let us see if it is possible to reconstruct, at least in principle, how did the neural code come into being in the history of life.

3.2. Neural code phase 1 – beginning

The nervous system is made of three types of neurons: (1) the *sensory neurons* transmit to the brain the signals produced by the sense organs, (2) the *motor neurons* deliver signals from the brain to the motor organs (muscles and glands), and (3) the *intermediate neurons* provide a bridge between them. In some cases the sensory neurons are directly connected to the motor neurons, thus forming a *reflex arch*, a system that produces a quick stimulus-response effect known as *reflex action*.

The first nervous systems were probably a collection of reflex arches, as it is still the case in a few primitive animals, and it is likely that the first intermediate neurons evolved as an extension of those arches. Once in existence, however, in addition to *transmitting* electrical signals they started *processing* them and this new function fuelled their evolution into increasingly complex systems. This is because the behaviour of an animal must take into account a variety of cues from the environment, and to that purpose it is necessary that a motor organ receives signals from many sense organs and that a sense organ delivers signals to many motor organs.

The intermediate neurons solved that problem by developing multiple connections between sensory inputs and motor outputs, but it must be underlined that they evolved in two very different directions. One was the formation of neural networks that are totally unconscious and provide a sort of *automatic pilot* for the body. The other was the generation of *reactive neural states*, the precursors of instincts and feelings, and it was this second process that set in motion the evolution of the neural code.

It must be underlined that all sense organs send to the brain the same electrical signals (in the form of trains of action potentials) and could, in principle, evoke the same reactions. Today this does not happen because in embryonic development the intermediate neurons go through processes of differentiation that allow them to respond in different ways to the same signals that come from different organs (Gilbert, 2006).

In the first animals, on the other hand, the ancestral neurons had not yet evolved the differentiation processes that we find in their descendants, and this means that they were responding in similar ways to the similar signals from the sense organs. The first neural responses, in other words, were little more than generic reactions but in ancestral times even vague stimulations could have been useful. That is probably the situation that existed at the time of the ancestral animals and our problem is to figure out, at least in principle, how could the neural code have evolved from such a beginning.

3.3. Neural code phase 2 – evolution

The evolution of the genetic code has been illustrated by Jacques Ninio (1982) with a beautiful metaphor. He pointed out that in any hotel, in addition to the familiar keys that open individual doors, there is a pass-key that opens all doors. At first, one may think that the pass-key is the most complex of all, whereas the truth is exactly the opposite. The pass-key is the simplest because what is complex in a key is not the ability to open a door but the ability to open one particular door *and not all the others*.

Ninio remarked that the adaptors of the genetic code can be compared to keys that open individual doors, whereas their common ancestor was like a pass-key that could open all doors. A particularly interesting point is that Ninio's metaphor can also be used to illustrate the evolution of the neural code. In this case, in fact, it is doubly useful because the adaptors of the neural code are intermediate neurons and these cells undergo two distinct processes of diversification: one in embryonic development and the other in evolution.

The diversification in embryonic development is necessary because it is only that process that allow the intermediate neurons to respond in different ways to the same electrical signals that come from different sense organs. This, incidentally, has an outstanding theoretical implication: the fact that the intermediate neurons diversify their responses during embryonic development means that those responses *can be modified according to need* and are therefore based on *arbitrary* rules; which amounts to saying that a *neural code* does exist in the intermediate brain.

The diversification of the intermediate neurons in evolution was necessary because the ancestral neurons had not yet evolved the differentiation processes that we find in their descendants, and this means that the ancestral neural code was different from the modern code. Without the modern differentiation processes, the ancestral neurons were bound to respond in the same way to the electrical signals that they were receiving from the sense organs. The first neural code, in other words, was totally ambiguous (like a pass-key that opens all doors).

The crucial point is that the ambiguity of the neural code was affecting not only the neural states that are induced in the intermediate neurons, but also the signals that these neurons deliver to the motor organs. The ambiguity of the neural code, in other words, was generating ambiguity in all reactions of the body, and anything that could improve these reactions was bound to be favoured in evolution.

The ability of the animals to respond to the signals from the environment was directly affected by the ambiguity of the neural code, and this ambiguity could be decreased only by introducing new differentiation steps in embryonic development, a process that went on until the intermediate neurons acquired the ability to produce specific reactions to the signals from the sense organs (like keys that only open individual doors).

The ancestral animals could not evaluate the *future* benefits of a non-ambiguous neural code, but they could certainly evaluate the *immediate* benefits of any operation that improved the reactions of the body, and it was for this reason that evolution systematically decreased the ambiguity of the ancestral neural code.

It will be noticed that the evolution of the neural code took place by differentiation processes that went on in all intermediate neurons. It was not about *individual* features but about *collective* relationships. It was the evolution of a systems as a whole, an evolution that went on until any ambiguity in the neural code was completely removed and *biological specificity* in the nervous system came into being.

3.4. Neural code phase 3 – optimization

In the case of the genetic code it is possible to conceive countless alternatives simply by postulating different mappings between codons and amino acids, and the study of these alternative codes has shown that the genetic code went through a process of optimization in an early phase of its history.

In the case of the neural code this approach is not practicable because its rules are not implemented by stable molecules, but by fleeting brain processes that are much harder to pin down. Despite this, however, we can still say that the neural code too went through a process of optimization.

A first hint of this potential was given by Aldous Huxley (1954) when he remarked that "*The function of the nervous system is to protect us from being overwhelmed and confused by a mass of largely useless and irrelevant knowledge, by shutting out most of what we should otherwise perceive or remember at any moment, and leaving only that very small and special selection which is likely to be practically useful.*"

This ability of the nervous system to select a small part of the incoming signals is in fact a form of optimization, in the sense that the brain had to learn how to obtain a maximum result from a minimum amount of data. This concept was formulated in more precise terms by Ross Ashby with the *free-energy principle*, the idea that self-organizing systems maintain their internal order by restricting themselves to a limited number of states (Ashby, 1962).

It has been shown that the free-energy principle is closely associated with the law of entropy and with the principle of least action and is at work in many processes, in particular in self-assembly, pattern formation, cybernetic control and autopoiesis (Jaynes, 1957; Evans, 2003; Ortega and Braun, 2012). The principle has also been applied to animal evolution by Karl Friston, who proposed that the nervous system evolved in a way that tended to minimize the difference between the information received from the sense organs and the model of the world built by the brain (Friston et al., 2006; Friston, 2010, 2012).

The free-energy principle, in other words, allows animals to catch increasing amounts of reality by evolving relationships between mental images that represent at least some of the relationships that exist between objects in the physical world. It is true that the brain has no direct access to reality because its information is filtered by the sense organs, but we can still say that our perceptions must reflect some aspects of reality otherwise we could not survive. François Jacob has expressed this concept with admirable clarity: "*If the image that a bird gets of the insects it needs to feed its progeny does not reflect at least some aspects of reality, there are no more progeny. If the representation that a monkey builds of the branch it wants to leap to has nothing to do with reality, then there is no more monkey. And if this did not apply to ourselves, we would not be here to discuss this point*" (Jacob, 1982).

It is very likely, in short, that the nervous system evolved with a mechanism that was closely related to the free-energy principle, and this is a minimization principle so it is legitimate to conclude that a process of optimization did take place in the evolution of the brain.

3.5. Neural code phase 4 – major transition

Feelings, instincts and emotions are referred to as *first-person* experiences because they are experienced directly, without intermediaries. They make us feel that we *know* our body, that we are in charge of its movements, that we are conscious beings and that we live a personal life. Above all, they are quintessentially *private* internal states, and this makes it impossible to share them with other people.

The goal of science is to produce reliable models of what exists in nature, and first-person experiences are undoubtedly part of nature, so we need models that help us, at least in principle, to understand them.

Let us take, for example, the case in which a toe is injured. We know that signals are immediately sent to the brain that processes them and delivers orders to the motor organs that spring the body into action. Here we have two distinct players where one (the intermediate brain) is the observer and the other (the injured toe) is the observed. It is the observer that receives signals from the toe and transforms them into a feeling of pain, but then something extraordinary happens. We do not feel the pain in the brain, where the feeling is created, but in the toe, where the injury took place. Observer and observed have collapsed into one, and it is precisely this short-circuit between them that generates a first-person experience.

Something similar takes place when we receive signals from the environment, for example when we look at a tree. In this case, an image is formed on the retina and the retina sends signals to the brain. Again, there is a physical separation between the sender and the receiver of signals, and yet we do not see an image on the retina, where the visual signals are generated, nor in the brain, where they are processed. What we see is a tree in the outside world. This is again a first-person experience and again it is generated by a physiological process that somehow creates a short-circuit between observer and observed.

This tells us that first-person experiences are nothing elementary and indivisible. On the contrary, they are the result of complex operations where highly differentiated cells act in concert to create a physiological short-circuit between body and brain, between observer and observed, between senders and receivers of neural signals.

That kind of complexity was necessarily the result of a long evolutionary process, and that process could start only when feelings and instincts were playing specific roles in animal behaviour, i.e., only when the ambiguity of the ancestral neural code was completely eliminated.

The origin of the neural code, in other words, set in motion a true biological revolution, a major transition that transformed the unconscious brain of the ancestral animals into the feeling brain of the modern animals. The result was an absolute novelty: it was the origin of *consciousness*, the origin of *subjectivity*, the origin of *first-person* experiences, in short, the origin of *mind*.

This is the *code theory of mind*, the idea that there has been a *neural code* at the origin of mind as there has been a genetic code at the origin of life. It is also the idea that there are neurological processes that create short-circuits between brain and body and give origin to first-person experiences, to the feeling that we are conscious subjects and not automatons (Barbieri, 2011, 2015).

3.6. Neural code phase 5 – conservation

There is ample evidence that virtually all animals have the same basic instincts and feelings. They all have the imperative to *survive* and to *reproduce*. They all experience hunger and thirst, fear and aggression, and they are all capable of reacting to stimuli such as light, sound, pressure and temperature. The basic feelings and instincts, in short, are virtually universal in animals, and this means that they appeared in an ancestral population and have been highly conserved ever since. After the common ancestor, on the other hand, the intermediate brain has continued to evolve and to change in many different ways. The invertebrates, for example, adopted a streamlining strategy that reduced their brain development to the bare essentials, whereas the vertebrates explored almost without limits the potentialities of the intermediate brain.

The conservation of the basic instincts and feelings, in other words, has gone hand in hand with the transformation of the intermediate brain, a pattern that has also been observed in the evolution of the cells, where the genetic code has been highly conserved whereas the apparatus of protein synthesis has continued to change. In both cases we have a system where virtually everything is on the move, except a fundamental set of rules, and this strongly suggests that a neural code does exist and that, like the genetic code, it has been highly conserved after its appearance in a common ancestor. The existence of the neural code, in other words, is the most parsimonious explanation of the experimental fact that the basic instincts and feelings of the animals have been strongly conserved in evolution.

Another faculty of the brain that is virtually universal is the ability to *dream*. Dreams are neural states that are clearly *manufactured* by the brain, and are therefore brain *artefacts* that are assembled according to a set of rules. Again, the existence of the neural code is the most parsimonious explanation of what we actually observe in Nature.

The operations of the intermediate brain are set in motion by the signals received from the sense organs, but these organs arise from the basic histological tissues of the body, and these tissues (epithelial, connective, muscular and nervous tissues) are the same in all triploblastic animals. All signals that are delivered to the brain, in other words, are produced by organs that arise from a limited number of universal tissues, and represent therefore a limited number of universal inputs. The basic feelings and instincts, on the other hand, are found in virtually all triploblastic animals and represent a limited number of universal outputs.

What we observe, in short, is a universal set of basic histological tissues on one side, a universal set of basic animal feelings on the other side, and a set of neural transformation processes in between. The most parsimonious explanation is that the neural processes in between are also a universal set of operations. And since there is no necessary physical link between sense organs and feelings, we conclude that the bridge between them can only be provided by the arbitrary rules of a highly conserved *neural code*.

4. Part 3: codes and macroevolution

4.1. Two great families of biological codes

The genetic code was the first of a long succession of organic codes that have appeared in the history of life. Among them, the *sequence codes* (Trifonov, 1989, 1996, 1999), the *sugar code* (Gabius, 2000, 2009), the *signal transduction codes* (Barbieri, 2003), the *splicing codes* (Barbieri, 2003; Fu, 2004; Buratti et al., 2006; Wang and Cooper, 2007), the *compartment codes* (Barbieri, 2003), the *tubulin code* (Verhey and Gaertig, 2007; Janke, 2014), the *nuclear signalling code* (Maraldi, 2008), the *ubiquitin code* (Komander and Rape, 2012), the *molecular codes* (De Beule et al., 2011; Görlich et al., 2011; Görlich and Dittrich, 2013; Dittrich, 2018) and the *lamin code* (Maraldi, 2018).

The evolution of life took place exclusively in single cells for about three billion years, but eventually some eukaryotes gave origin to multicellular creatures and new organic codes came into being. Among them: the *Hox* code (Hunt et al., 1991; Kessel and Gruss, 1991), the *adhesion codes* (Redies and Takeichi, 1996; Shapiro and Colman, 1999; Faria, 2018), the *histone code* (Strahl and Allis, 2000; Turner, 2000, 2007; Kühn and Hofmeyr, 2014; Prakash and Fournier, 2018), the *transcriptional codes* (Jessell, 2000; Marquard and Pfaff, 2001), the *apoptosis code* (Basañez and Hardwick, 2008; Füllgrabe et al., 2010), and the *bioelectric code* (Tseng and Levin, 2013; Levin, 2014).

With the origin of the nervous systems, a completely new type of biological entities came into being; entities such as *action potentials* and *neuron firings* at the cellular level, and entities such as *instincts and feelings* at the organism level. In addition to organic evolution, based on organic molecules, what came into being was *neural evolution*, based on neural states, and even in this new world we find codes.

The neural code was the forerunner of a new family of biological codes that evolved in nervous systems. It has been reported, for example, a *neural code for mechanical stimuli* (Nicolelis and Ribeiro, 2006; Nicolelis, 2011), a *neural code for taste* (Di Lorenzo, 2000; Hallock and Di Lorenzo, 2006), a *synaptic code* (Hart et al., 1995; Szabo and Soltesz, 2015), a *space code* in the hippocampus (O'Keefe and Burgess, 1996, 2005; Hafting et al., 2005; Brandon and Hasselmo, 2009), the *acoustic codes* (Farina and Pieretti, 2014; Farina, 2018) and the *olfactory code* (Grabe and Sachse, 2018).

The evolution of the nervous systems, on the other hand, produced not only neural codes but also the animal ability of interpreting the world. The difference between coding and interpretation is beautifully illustrated by a classical example of animal behavior. When a snake chases a prey and the prey hides behind a tree, the snake *stops* chasing. When a wolf chases a prey and the prey hides behind a tree, the wolf *goes on* chasing. The snake is only using *codified rules*, whereas the wolf makes an act of *interpretation*. The wolf makes a 'mental jump beyond the appearances', and that is what interpretation is: a 'jump to conclusions', a process that Charles Peirce called *abduction*.

An act of interpretation, in other words, is a form of semiosis because it gives meanings to something, but it is different from coding because it is not based on a fixed set of coding rules but on processes of abduction. The most traditional form of semiosis, on the other hand, is language and this gives us a problem: is language a third type of semiosis or can we reduce it to coding and interpretation?

5. The origin of language

The study of language is complicated by the fact that there is no consensus on its definition. Some have proposed that language is based on a universal grammar, and therefore on the rules of a code, whereas others insist that language is a social phenomenon. This is why it is important to make a sharp distinction between a code and the apparatus that is using that code. In protein synthesis, for example, this distinction is clear because we know that the genetic code is a set of rules that are used by the apparatus that makes proteins. In linguistics, instead, the term *language* is often used in both senses so it is necessary to underline that the apparatus of language is a system that may or may not contain a code but in any case it is much more complex than a set of coding rules.

This amounts to saying that the origin of language consists in two distinct problems: one is the origin of the *apparatus of language*, the other is the origin of the *code of language*, if such a code does exist.

As for the first problem, we can get an idea of what is involved in the apparatus of language from a brief summary of what takes place in the first years of human development.

It has been calculated that our species should have a gestation period of 21 months in order to complete all processes of foetal development (Portmann, 1941, 1945; Gould, 1977). A newborn human baby, in other words, is in effect *a premature foetus*, and the whole first year of life is but a continuation of the foetal stage. This peculiarity is due to the fact that the birth canal can only cope with a limited increase of the foetus size, and any extension of the foetal period had to be accompanied by an anticipation of the time of birth. The result is that our foetal development became split into two distinct phases – intrauterine and extrauterine – and eventually the extrauterine phase (12 months) became the longest of the two.

The crucial point is that the second part of foetal development is a phase of intense *brain wiring*. In all other mammals the major part of brain wiring takes place in the protected environment of the uterus, whereas in our species it takes place outside the uterus, where the body is exposed to the turbulence of a constantly changing environment. In our species, in short, the extrauterine phase of foetal development creates the conditions for a unique type of brain wiring *and language does depend on brain wiring*.

Another crucial point is that infants acquire a language only if their postnatal development is *assisted*, i.e., only if they interact with other human beings, and this tells us that the acquisition of language is also a *social* process and not only an individual one.

Let us now turn to the second problem: is there a language code in the apparatus of language as there is a genetic code in the apparatus of protein synthesis?

Here we need to distinguish between *written* language and *spoken* language. Written language is clearly based on codes because it is a mapping between written signs and sounds, whereas spoken language is a much more complex phenomenon. The crucial point is that written language is totally dependent on spoken language and this implies that the codes of written language could hardly exist without a prior code in spoken language, so it is highly likely that a code of language does exist in the apparatus of language even if we have not yet been able to pin it down.

Language, in conclusion, is acquired by infants with a variety of mechanisms that includes codes and interpretation but it is not reducible to them and represents therefore a third form of semiosis.

6. The codes of culture

Language gave origin to culture and culture is based on cultural codes, so there are three great classes of codes in the living world: the organic codes, the neural codes and the cultural codes. The key point is that there is an enormous difference between them: the cultural codes are ephemeral, short-lived and constantly changing, whereas the biological codes (organic codes and neural codes) are the most conserved entities in evolution. How can we explain this massive difference between the most conserved and the least conserved rules in the history of life?

All codes are sets of arbitrary rules that create a bridge between the objects of two independent worlds that function as signs and meanings by means of other objects called *adaptors*, and the differences between the codes are largely due to the differences that exist between their objects.

In the genetic code the signs are codons, the meanings are amino acids and the adaptors are the transfer-RNAs that create a mapping between them. All these molecules are genetically inherited and have been highly conserved in evolution because only their conservation assures that the same biological characteristics are transmitted from one generation to the next.

In the neural code the adaptors are the neurons of the intermediate brain, the cells that receive signals from the sense organs and transform them into instincts and feelings according to a fixed set of rules. It must be underlined, however, that these cells must undergo specific processes of differentiation in order to perform their functions. This is because all sense organs send the same types of electrical signals to all intermediate neurons, and it is only the differentiation processes of embryonic development that allow the neurons to respond in different ways to the same signals. In the neural code, in other words, signs, meanings and adaptors are not inherited as they are in the genetic code. They come into being only after processes of embryonic differentiation, and we know that they have been highly conserved in evolution because there is ample evidence that virtually all animals have the same basic instincts and feelings and are capable of reacting to external stimuli such as light, sound, pressure and temperature.

In the codes of culture the adaptors are much more complex than molecules and cells: they are the human beings themselves, the systems that actually associate signs to meanings in all cultural codes. In this case, however, what makes the difference is not the complexity of the adaptors per se. It is the fact that the signs and the meanings of the cultural codes are not transmitted by genes or by embryonic processes: they are acquired by processes of imitation, interpretation and learning in the first years of postnatal development and only when this development is assisted by other human beings.

This means that the signs and the meanings of the cultural codes are affected by cultural variations and can change according to cultural circumstances, which explains why they are intrinsically short-lived. The codes of culture, in other words are fundamentally different from the codes of nature and this implies that the general model described in this paper does not apply to them. It is a model for codes that are highly conserved in evolution and not for the ephemeral man-made codes of culture.

7. Two mechanisms of evolution

The fact that many biological codes exist today in living systems means that they came into being during the history of life, and this suggests that *coding* is a mechanism that Nature has employed many times in the course of evolution. Today, on the other hand, it is almost universally accepted that *natural selection* is the sole mechanism of evolution, and this means that we need to take a closer look at the relationship that exists between natural selection, biological codes and evolution. Let us start from this question: where does natural selection come from?

The copying of the genes is the elementary act that leads to *heredity*, and when the process of copying goes on and on indefinitely another phenomenon comes into being. Copying mistakes become inevitable and in a world of limited resources not all changes can be realized and a selection is bound to take place. Molecular copying, in short, leads to heredity, and the indefinite repetition of molecular copying in a world of limited resources leads to natural selection. That is how natural selection came into existence. Molecular copying started it and molecular copying has perpetuated it ever since. Which means that *natural selection would be the sole mechanism of evolution if molecular copying were the sole basic mechanism of life.*

The discovery of the genetic code, however, has proved that there are *two* distinct molecular mechanisms at the basis of life, the *copying* of the genes and the *coding* of proteins. The discovery of other organic codes, furthermore, allows us to generalize this conclusion because it proves that coding is not limited to protein synthesis. Life, in other words, is not based on copying alone. It is based on copying *and* coding, and these two molecular mechanisms give origin to two distinct mechanisms of evolution because an evolutionary mechanism is but the long term result of a fundamental molecular mechanism.

More precisely, the existence of copying and coding at the molecular level means that there are two distinct types of evolutionary change: *evolution by natural selection*, based on copying, and *evolution by natural conventions*, based on coding (Barbieri, 1985, 2003).

But can we prove it? The only way to prove that natural selection and natural conventions are two distinct mechanisms is to prove that *coding cannot be reduced to copying.*

This in fact turns out to be fairly straightforward because copying and coding involve two very different entities. A variation in the copying of a gene changes the linear sequence, i.e., the *information* of that gene. A variation in a coding rule, instead, changes the *meaning* of that rule. The great difference that exists between copying and coding, and therefore between natural selection and natural conventions, comes from the difference that exists between 'information' and 'meaning'.

The fact remains, however, that copying is a universally accepted mechanism of evolution, but coding is not, and this has probably a simple explanation: it is due to the fact that we understand the mechanism of copying but do not see how the arbitrary rules of a code could have come into existence. This is why we need a theoretical framework that makes us understand the origin of biological codes, and the model presented in this paper is a first step in that direction.

Acknowledgments

I am most grateful to Massimo Di Giulio, Almo Farina and Nikola Štambuk whose comments have greatly improved the first version of this manuscript.

References

Ashby, W.R., 1962. Principles of the self-organizing system. In: Von Foerster, H., Zopf Jr.G.W. (Eds.), Principles of Self-Organization: Transactions of the University of Illinois Symposium. Pergamon Press, London, UK, pp. 255–278.
Barbieri, M., 1985. The Semantic Theory of Evolution. Harwood Academic Publishers, London & New York.
Barbieri, M., 2003. The Organic Codes. An Introduction to Semantic Biology. Cambridge University Press, Cambridge, UK.
Barbieri, M., 2011. Origin and evolution of the brain. Biosemiotics 4 (3), 369–399.
Barbieri, M., 2012. Codepoiesis – the deep logic of life. Biosemiotics 5 (3), 297–299.
Barbieri, M., 2015a. Code Biology. A New Science of Life. Springer, Dordrecht.
Barbieri, M., 2015b. Evolution of the genetic code: the ribosome-oriented model. Biol. Theory 10 (4), 301–310.
Basañez, G., Hardwick, J.M., 2008. Unravelling the Bcl-2 apoptosis code with a simple model system. PLoS Biol. 6 (6), e154 doi:10.137/journal.pbio.0060154.
Blobel, G., 1980. Intracellular membrane topogenesis. Proc. Natl. Acad. Sci. U. S. A. 77, 1496–1500.
Bollenbach, T., Kalin Vetsigian, K., Kishony, R., 2007. Evolution and multilevel optimization of the genetic code. Genome Res. 17, 401–404.
Brandon, M.P., Hasselmo, M.E., 2009. Sources of the spatial code within the hippocampus. Biol. Rep. 1, 3–7.
Budisa, N., 2004. Prolegomena to future efforts on genetic code engineering by expanding its amino acid repertoire. Angew. Chem. Int. Ed. 43, 3387–3428.
Buratti, E., Baralle, M., Baralle, F.E., 2006. Defective splicing, disease and therapy: searching for master checkpoints in exon definition. Nucleic Acids Res. 34, 3494–3510.
Cavalier-Smith, T., 2000. Membrane heredity and early chloroplast evolution. Trends Plant Sci. 5, 174–182.
Churchland, P.S., Sejnowski, T.J., 1993. The Computational Brain. MIT Press, Cambridge MA.
Crick, 1981. Life Itself. Simon and Schuster, New York.
Crick, F.H.C., 1994. The Astonishing Hypothesis: The Scientific Search for the Soul. Scribner, New York.
De Beule, J., Hovig, E., Benson, M., 2011. Introducing dynamics into the field of Biosemiotics. Biosemiotics 4 (1), 5–24.
Di Lorenzo, P.M., 2000. The neural code for taste in the brain stem: response profiles. Physiol. Behav. 69, 87–96.
Dittrich, P., 2018. Towards measuring the semantic capacity of a physical medium demonstrated with elementary cellular automata. BioSystems 164, 177–185.
Drummond, D.A., Wilke, C.O., 2008. Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. Cell 134, 341–352.
Edelman, G.M., 1989. Neural Darwinism. The Theory of Neuronal Group Selection. Oxford University Press, New York.
Evans, D.J., 2003. A non-equilibrium free energy theorem for deterministic systems. Mol. Phys. 101, 15551–15554.
Faria, M., 2018. Aggregating, polarizing, networking – the evolution of cell adhesion codes. BioSystems 164, 60–67.
Farina, A., 2018. Ecoacustic codes and ecological complexity. BioSystems 164, 147–154.
Farina, A., Pieretti, N., 2014. Acoustic codes in action in a soundscape context. Biosemiotics 7 (2), 321–328.
Fitch, W.M., Upper, K., 1987. The phylogeny of tRNA sequences provides evidence for ambiguity reduction in the origin of the genetic code. Cold Spring Harb. Symp. Quant. Biol. 52, 759–767.
Fodor, J., 1975. The Language of Thought. Thomas Crowell Co., New York.
Fodor, J., 1983. The Modularity of Mind. An Essay on Faculty Psychology. MIT Press, Cambridge MA.
Fox, G.E., 2010. Origin and evolution of the ribosome. Cold Spring Harb. Perspect. Biol. 2,

a003483.
Freeland, S.J., Hurst, L.D., 1998. The genetic code is one in a million. J. Mol. Evol. 47, 238–248.
Friston, K., 2010. The free-energy principle: a unified brain theory? Nat. Rev. Neurosci. 11 (2), 127–138.
Friston, K., 2012. A free energy principle for biological systems. Entropy 14, 2100–2121.
Friston, K., Kilner, J., Harrison, L., 2006. A free energy principle for the brain. J. Physiol. Paris 100 (1–3), 70–87.
Fu, X.D., 2004. Towards a splicing code. Cell 119, 736–738.
Füllgrabe, J., Hajji, N., Joseph, B., 2010. Cracking the death code: apoptosis-related histone modifications. Cell Death Differ. 17, 1238–1243.
Gabius, H.-J., 2000. Biological information transfer beyond the genetic code: the sugar code. Naturwissenschaften 87, 108–121.
Gabius, H.-J., 2009. The Sugar Code. Fundamentals of Glycosciences. Wiley-Blackwell.
Gilbert, S.F., 2006. Developmental Biology. Sinauer Associates, Sunderland, Massachusetts USA.
Gilis, D., Massar, S., Cerf, N.J., Rooman, M., 2001. Optimality of the genetic code with respect to protein stability and amino-acid frequencies. Genome Biol. 2, 41–49.
Görlich, D., Dittrich, P., 2013. Molecular codes in biological and chemical reaction networks. PLoS One 8 (1), 694. https://doi.org/10.1371/journal.pone.0054694. e54.
Görlich, D., Artmann, S., Dittrich, P., 2011. Cells as semantic systems. Biochim. Biophys. Acta 1810 (10), 914–923.
Gould, S.J., 1977. Ontogeny and Phylogeny. Harvard University Press, Cambridge MA.
Gould, S.J., 2002. The Structure of Evolutionary Theory. Harvard University Press, Cambridge MA.
Grabe, V., Sachse, S., 2018. Fundamental principles of the olfactory code. BioSystems 164, 94–101.
Hafting, T., Fyhn, M., Molden, S., Moser, M.B., Moser, E.I., 2005. Microstructure of a spatial map in the entorhinal cortex. Nature 436, 801–806.
Haig, D., Hurst, L.D., 1991. A quantitative measure of error minimization in the genetic code. J. Mol. Evol. 33, 412–417.
Hallock, R.M., Di Lorenzo, P.M., 2006. Temporal coding in the gustatory system. Neurosci. Behav. Rev. 30, 1145–1160.
Harold, F.M., 2005. Molecules into cells: specifying spatial architecture. Microbiol. Mol. Biol. Rev. 69, 544–564.
Harold, F.M., 2014. In Search of Cell History. The Evolution of Life's Building Blocks. The University of Chicago Press, Chicago and London.
Hart, A.C., Sims, S., Kaplan, J.M., 1995. Synaptic code for sensory modalities revealed by C. Elegans GLR-1 glutamate receptor. Nature 378, 82–85.
Hartman, M.C.T., Josephson, K., Lin, C.-W., Szostak, J.W., 2007. An expanded set of amino acid analogs for the ribosomal translation of unnatural peptides. PLoS One 2 (10), e972. https://doi.org/10.1371/journal.pone.0000972.
Holland, J.A., 1992. Adaptation in Natural and Artificial Systems. MIT Press, Cambridge MA.
Hopfield, J.J., 1982. Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. U. S. A. 79, 2554–2558.
Hunt, P., Whiting, J., Nonchev, S., Sham, M.-H., Marshall, H., Graham, A., Cook, M., Alleman, R., Rigby, P.W., Gulisano, M., 1991. The branchial *Hox* code and its implications for gene regulation, patterning of the nervous system and head evolution. Development 2, 63–77.
Huxley, A., 1954. The Doors of Perception. Chatto and Windus, London.
Itzkovitz, S., Alon, U., 2007. The genetic code is nearly optimal for allowing additional information within protein-coding sequences. Genome Res. 17, 405–412.
Jacob, F., 1982. The Possible and the Actual. Pantheon Books, New York.
Janke, C., 2014. The tubulin code: molecular components, readout mechanisms, and functions. J. Cell Biol. 206 (4), 461–472.
Jaynes, E.T., 1957. Information theory and statistical mechanics. Phys. Rev. Ser. II 106 (4), 620–630.
Jessell, T.M., 2000. Neuronal specification in the spinal cord: inductive signals and transcriptional codes. Nat. Genet. 1, 20–29.
Johnson-Laird, P.N., 1983. Mental Models. Harvard University Press, Cambridge MA.
Kessel, M., Gruss, P., 1991. Homeotic tansformation of murine vertebrae and concomitant alteration of *Hox* codes induced by retinoic acid. Cell 67, 89–104.
Komander, D., Rape, M., 2012. The ubiquitin code. Annu. Rev. Biochem. 81, 203–229.
Koonin, E.V., Martin, W., 2005. On the origin of genomes and cells within inorganic compartments. Trends Genet. 21, 647–654.
Kühn, S., Hofmeyr, J.-H.S., 2014. Is the "Histone Code" an organic code? Biosemiotics 7 (2), 203–222.
Kurland, C.G., 1970. Ribosome structure and function emergent. Science 169, 1171–1177.
Lecompte, O., Ripp, R., Thierry, J.C., Moras, D., Poch, O., 2002. Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale. Nucleic Acid Res. 30 (24), 5382–5390.
Levin, M., 2014. Endogenous bioelectrical networks store non-genetic patterning information during development and regeneration. J. Physiol. (Lond.) 592, 2295–2305.
Lombard, J., Moreira, D., 2011. Origins and early evolution of the mevalonate pathway of isoprenoid biosynthesis in the three domains of life. Mol. Biol. Evol. 28 (1), 87–99.
Maizels, N., Weiner, A.M., 1987. Peptide-specific ribosomes, genomic tags and the origin of the genetic code. Cold Spring Harb. Symp. Q. Biol. 52, 743–757.
Maraldi, N.M., 2008. A lipid-based code in nuclear signalling. In: Barbieri, M. (Ed.), The Codes of Life: The Rules of Macroevolution. Springer, Dordrecht, pp. 207–221.
Maraldi, N.M., 2018. The Lamin code. BioSystems 164, 68–75.
Marquard, T., Pfaff, S.L., 2001. Cracking the transcriptional code for cell specification in the neural tube. Cell 106, 651–654.
Maturana, H.R., Varela, F.J., 1980. Autopoiesis and Cognition: The Realisation of the Living. D. Reidel Publishing Company, Dordrecht, Holland.
Morgan Lloyd, C., 1923. Emergent Evolution. Williams and Norgate, London.
Mulkidjanian, A.Y., Makarova, K.S., Galperin, M.Y., Koonin, E.V., 2007. Inventing the dynamo machine: the evolution of the F-type and V-type ATPases. Nat. Rev. Microbiol. 5, 892–899.
Nicolelis, M., 2011. Beyond Boundaries: The New Neuroscience of Connecting Brains with Machines and How It Will Change Our Lives. Times Books, New York.
Nicolelis, M., Ribeiro, S., 2006. Seeking the neural code. Sci. Am. 295, 70–77.
Ninio, J., 1982. Molecular Approaches to Evolution. Pitman books, London.
Nitta, I., Kamada, Y., Noda, H., Ueda, T., Watanabe, K., 1998. Reconstitution of peptide bond formation. Science 281, 666–669.
Nomura, M., Tissières, A., Lengyel, P., 1974. Ribosomes. Cold Spring Harbor Monograph Series.
Novozhilov, A.S., Wolf, Y.I., Koonin, E.V., 2007. Evolution of the genetic code: partial optimization of a random code for robustness to translation error in a rugged fitness landscape. Biol. Direct 2 (24).
O'Keefe, J., Burgess, N., 1996. Geometric determinants of the place fields of hippocampal neurons. Nature 381, 425–428.
O'Keefe, J., Burgess, N., 2005. Dual phase and rate coding in hippocampal place cells: theoretical significance and relationship to entorhinal grid cells. Hippocampus 15, 853–866.
Ortega, P.A., Braun, D.A., 2012. Thermodynamics as a theory of decision-making with information processing costs. Proc. R. Soc. A 469 (2153) 20120683.
Osawa, S., 1995. Evolution of the Genetic Code. Oxford University Press, New York.
Pattee, H.H., 2001. The physics of symbols: bridging the epistemic cut. BioSystems 60, 5–21.
Portmann, A., 1941. Die Tragzeiten Der Primaten und die Dauer Der Schwangerschaft beim menschen: ein problem der vergleichen Biologie. Rev. Suisse Zool. 48, 511–518.
Portmann, A., 1945. Die Ontogenese Des Menschen als problem der Evolutionsforschung. Verh. Schweiz. Naturf. Ges. 125, 44–53.
Prakash, K., Fournier, D., 2018. Evidence for the implication of the histone code in building the genomestructure. BioSystems 164, 49–59.
Redies, C., Takeichi, M., 1996. Cadherine in the developing central nervous system: an adhesive code for segmental and functional subdivisions. Dev. Biol. 180, 413–423.
Rumelhart, D.E., McClelland, J.L., 1986. Parallel Distributed Processing: Explorations in the Microstructure of Cognition. MIT Press, Cambridge MA.
Sapp, J., 1987. Beyond the Gene: Cytoplasmic Inheritance and the Struggle for Authority in Genetics. Oxford University Press, Oxford UK.
Schimmel, P., 1987. Aminoacyl tRNA synthetases: general scheme of structure-function relationship in the polypeptides and recognition of tRNAs. Ann. Rev. Biochem. 56, 125–158.
Schimmel, P., Giegé, R., Moras, D., Yokoyama, S., 1993. An operational RNA code for amino acids and possible relationship to genetic code. Proc. Natl. Acad. Sci. U. S. A. 90, 8763–8768.
Searle, J.R., 1980. Minds, brains and programs. Behav. Brain Sci. 3, 417–457.
Searle, J.R., 2002. Consciousness and Language. Cambridge University Press, Cambridge UK.
Shapiro, L., Colman, D.R., 1999. The diversity of cadherins and implications for a synaptic adhesive code in the CNS. Neuron 23, 427–430.
Strahl, B.D., Allis, D., 2000. The language of covalent histone modifications. Nature 403, 41–45.
Szabo, G.G., Soltesz, I., 2015. Pass-through code of synaptic integration. Neuron 87, 1124–1126.
Trifonov, E.N., 1989. The multiple codes of nucleotide sequences. Bull. Math. Biol. 51, 417–432.
Trifonov, E.N., 1996. Interfering contexts of regulatory sequence elements. Cabios 12, 423–429.
Trifonov, E.N., 1999. Elucidating sequence codes: three codes for evolution. Ann. N. Y. Acad. Sci. 870, 330–338.
Tseng, A.S., Levin, M., 2013. Cracking the bioelectric code. Probing endogenous ionic controls of pattern formation. Commun. Integr. Biol. 6 (1), 1–8.
Turner, B.M., 2000. Histone acetylation and an epigenetic code. BioEssays 22, 836–845.
Turner, B.M., 2007. Defining an epigenetic code. Nat. Cell Biol. 9, 2–6.
Verhey, K.J., Gaertig, J., 2007. The tubulin code. Cell Cycle 6 (17), 2152–2160.
Wang, G.S., Cooper, T.A., 2007. Splicing in disease: disruption of the splicing code and the decoding machinery. Nat. Rev. Genet. 8, 749–761.
Woese, C.R., 1965. Order in the genetic code. Proc. Natl. Acad. Sci. U. S. A. 54, 71–75.
Woese, C.R., 1987. Bacterial evolution. Microbiol. Mol. Biol. Rev. 51, 221–271.
Woese, C.R., 2000. Interpreting the universal phylogenetic tree. Proc. Natl. Acad. Sci. U. S. A. 97, 8392–8396.
Woese, C.R., 2002. On the evolution of cells. Proc. Natl. Acad. Sci. U. S. A. 99, 8742–8747.
Woese, C.R., Dugre, D.H., Saxinger, W.C., Dugre, S.A., 1966. The molecular basis for the genetic code. Proc. Natl. Acad. Sci. U. S. A. 55, 966–974.
Wolf, Y.I., Koonin, E.V., 2007. On the origin of the translation system and the genetic code in the RNA world by means of natural selection, exaptation, and subfunctionalization. Biol. Direct 2, 14.