Journal of Natural Sciences Research
ISSN 2224-3186 (Paper) ISSN 2225-0921 (Online)
Vol.3, No.1, 2013

The Modular Concept of Gene

Zyri Bajrami

Department of Biology, University of Tirana, Bulevardi Zogu I, Tirana, Albania

E-mail: zyribajrami@yahoo.com

Abstract

In each living system it is common for a vital process to happen or for a function to be carried out; as a consequence, a module is formed. It happens when it encounters a stimulus or a need arises, when it establishes an effectors responding to the stimulus or the need, and when the connection between them is possible from a memory structure. Gene itself is a module: the gene is a fragment of nucleic acids, which in a particular type of module, labeled as a genetic module, play the part of a memory structure. It happens only with the modules for functions to be performed according to the information, which is carried by the memory structures.

Moreover, it is revealed that under the definition of the gene based on the modular concept, the information, otherwise denoting that the genes are to be found with the royal throne of the current biological thought, is not the mere sequencing of the nucleic acids on the basis of which another molecule of nucleic acids is created, a string of peptides, etc., but that type of information inherent in every memory structure, as is the gene itself, according to which is realized each function of livings. Consequently, the amount of information in the living systems is measured on the basis of the probability of fulfillment of functions from the respective memory structures.

Keywords: gene, modules, memory structures, information

1.Introduction

Any effort made to give a definition of the *gene* has stopped short of arriving at one such definition, acceptable to all, and this for the better part of a century. Its meaning has been and still continues to be "*a concept in tension"* (Falk 2000). The reason behind such a condition should be sought with the royal throne where the scientists have elevated it: the gene has been granted such an explanatory power that almost everything revolves around it.

As a matter of fact the absolute power, which the gene enjoys in biology, should rather rest on the information. However this information which makes gene synonymous to it, should by no means be the concept borrowed from physics or the technical sciences. In our undertaking we will prove that gene is a carrier of information or a memory structure (Bajrami 2008), why not the most fundamental memory structure of the living systems, but far from being the only one of the kind and unique. The gene, based on the thinking that will allow through this study, is not connected only with the correspondence of the nucleic language and the proteinic one, but also with the way as to how the nucleic acids, the proteins and the molecules or other structures are utilized by the organism to generate a phenotype (Noble 2008). Like any other memory structure, the gene forms a functional unit or a module, where the function is carried out based on an information (Jonker & Treur 2006). Just like an organism has its information and functions, which serve as the two crucial traits of the living systems (Emmeche 2002), in the same way each module has its own share of information and functions.

2. Modules

An idea has been given, which will serve as the starting point of our study: in the living systems the stimulus-response, cause-effect or input-output connections are triadic by nature (Hoffmeyer 2005). In our opinion, these connections are established by a memory structure (Bajrami 2008), which in the semiotic terminology is called representamen (Pierce, 1955). Let us give few examples.

- When someone touches by accident a piping hot object (stimulus O), one will snatch the hand away

In the above examples it is shown how a function is carried out from a triadic organization of elements (O, S and I) and consequently it is achieved to the idea that organizations of such kind immediately (response or interpretation I). The connection between stimulus-response is mediated by a sensitive cell, which plays the role of a *representamen,* which will be called a memory structure (S); because, on one hand, it has been able to perceive the heat (O) and has deposited the information about it; on the other hand, it establishes links to a muscular cell or with an effectors, which triggers the movement of the hand (I).

- It is known that upon a viral infection, the memory cells are found in mammals, which play the part of a memory structure (S); in the case of a second infection, they are able to recognize the previous virus as stimulus (O), and through the plasmatic cells they respond to it, or they are able to interpret its presence by forming the respective antibodies.

- Since the allosteric enzyme recognizes the initial and final substances of a biochemical reaction, it plays the role of a memory-borne structure (S). Hence, the allosteric enzyme responds to the presence in excess of the ultimate product of the biochemical reaction (O), by stopping the reaction. The same enzyme responds to the absence of the final product through the respective biochemical reaction.

- From the above examples, no exception should be made even our mental representation or meme, as an information model located in the brain (Dawkins 1982); playing the role of a memory structure it mediate the sign, that is the object apple (O) with the signer, that is the word apple (I).
are functional units or modules.

In all of the above it is shown that a leaving process happens when three conditions are fully satisfied.

- The living system encounters a stimulus, a need or a request (O).

- The living system forms an effector (I) towards the stimulus, the need or the request.

- In the living system, a connection between the stimulus-need-request-input (O) is established with the response-interpretation-output through a memory structure (S); a function is carried out based on the related information (Jonker & Treur 2006).

The three conditions referred to above are fully met in each module. The gene itself is such a module.

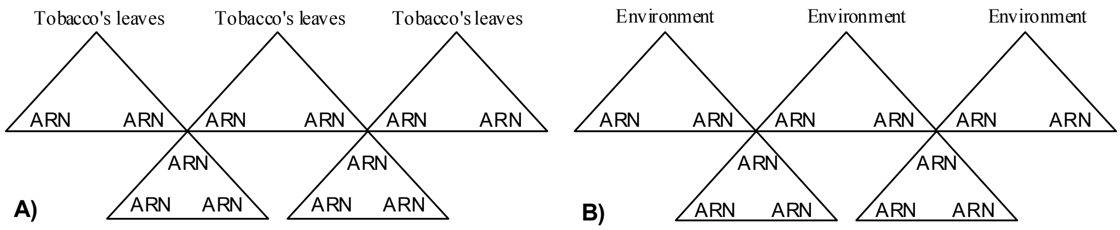
A gene, i.e. the gene responsible for the breakdown of the lactose in humans (LCT.P) will be considered to be a memory structure (S); in the past, as is the case with the sensitive cells nowadays, that one is able to extract the information from the environment (Cosmides & Tooby 1992). On the other hand, through the genetic code, the presence of lactose, which needs breaking down and which is a request to be satisfied or a lock to be opened up (O), the respective gene responds to it by forming the proper key, by producing the *lactase* enzyme (I).

The reader has no doubt about the fact that the gene is a memory structure which carries information. Nor are there any such suspicions concerning the product it forms that one carries out a certain function. But we are convinced that the suspicion lurks around the thought that another structure, besides the nucleic acids, that is an enzyme or another molecule, a cell or a cluster of cells, might be carriers of the information. This suspicion will be dissipated from our mind if we start thinking that anything which might be a source of information might as well be a carrier of information, when its presence correlate closely with the presence of the source of information (Shannon 1948). The fact that a module, whose memory-borne structure is not a fragment of the nucleic acids, carries out a function as vital as the genetic module does not reduce the importance of gene into dimness. The genetic modules, that are the modules whose memory structure is a segment of the nucleic acids, like any other module, when they contrive to perform the necessary functions, do convert the living system into a model of the environment where it dwells (Uexkull 1982). On the other hand, the living systems do replicate themselves or copy the material modules of themselves or replicators. In such cases the living systems form a type of specific module which will be called modules of reproduction.

2.1 Types and classification of modules

Starting from the thinking that prescribes the living organisms as made up of two types of modules, from survival modules and reproduction ones, it might be concluded that the most simple and primitive living systems should be composed of at least a module of each type. For instance the tobacco mosaic virus is considered as one of the simplest living systems, which is depicted in its own bimodular representation (Fig. 1A).

Fig. 1. The bimodular model of the simplest (A) and most primitive (B) living systems.



In Fig. 1 it is shown that module 1 is a reproduction module, because only in the presence of tobacco leaves, which in this case play the role of a memory structure (S), a molecule of the ancestor ARN (O) forms a successor molecule of ARN (I). The same holds true of the "ARN world" (Fig. 1B). A certain environment, similar to the tobacco leaves, has played the role of a memory structure (S). In this case, under module 1, the ARN molecule performs the function of a gene, while in the case of module 2 the same molecule plays the role of an enzyme, because by the time it doubles, it is able to interpret its own environment. This position in "ARN world" is also expressed in module 2 of the virus of the tobacco mosaic. Each time that ARN structure, as a memory structure (S) of this virus meets up with the tobacco leaves (O), it is able to interpret them by creating new molecules (I). The condition of the modules changes in all other organisms: the reproduction module converts itself into a survival module and vice-versa. Such a thing has happened because the living beings, by being confined as material units itself in the cell mode, experienced the so-called the internalization process of the outside environment in an information of nucleic acids as model of such an environment.

Yet in every living system the fulfillment of a function is determined by the number of molecules of nucleic acids which are formed, as well as, by the number of those which are living. Hence, on the basis of such function it is determined the value of the information of the molecules of the nucleic acids.

Under the circumstances the question is raised: how the living systems maintain themselves, their identity, for a relative long period of time.

Firstly, these systems do replicate themselves or form the so-called replicators. What is requested of such replicators is the loyalty of replication, the retention of the *ancestor-successor* identity and the formation of large number of off-springs with the least amount of energy consumed. But reproduction alone is not an adequate prerequisite for the retention of the identity of living system because the setting is always changing and that the formation of replicators depends on the cost-benefit balance. In order to clarify this situation a thought experiment might help us.

Let's suppose that two teams of children are busy playing in the city's playing ground. One group consists of 4 children and the other of 10. Let's also suppose that one kid from each group has lost his own key and all the other children have set about looking for it. The question is raised: Which group of children stands a better chance at finding the lost key?
The finding of the key depends at least on three factors:

First of all, the finding of the key is latched with the setting, relief and the vegetation cover, etc.

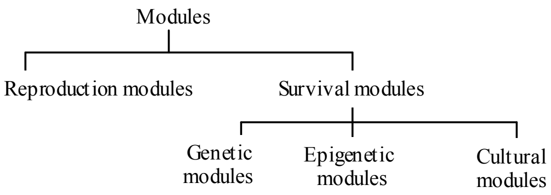
Secondly, the recovery of the key depends on the wisdom and the familiarity with the setting from each individual from the respective groups.

Lastly, the finding of the key depends on the size of the group that is from the individual members that make up the group. This means that the group consisting of 10 children stands a better chance at finding the key than the other group which has almost half that number.

The three factors that stipulate the finding of the key are also valid for the possibilities which determine the survival of one kind or for that matter the success in the game of life. In the jargon of the type these factors have to do with the environmental conditions, as well as with the presence of the various types of genotypes; the more varied they tend to be, the bigger are the chances at survival of the type. From this fact alone, we conclude that each genotype bears the value of a module for the type; each one of them carries out functions which are adaptable variants or alternatives of the behavior of the type against the environment.

In this way, the genotype or the offspring which comprises a type are its own modules; unlike the survival modules, they are called reproduction modules. The types of modules and classification are shown in Fig. 2.

Fig. 2 The types and classification of modules.



The classification of modules into reproduction and survival modules support the idea that the living organisms consist of parts (that is module-based in our thought), which function to provide their reproduction and survival (Maynard –Smith & Szathmary 1995).

2.2 The information and its related value

In the above examples, it is shown that in each module type the fulfillment of a function is made possible on the basis of the information that its memory structure carry out. It implies that if the value of probability for fulfillment of the function is high, then even the amount of information in the living systems, and exactly the information that is contained by its memory structure, might be measured on the value of performance probability of the function in the related module. Therefore, it is proposed that when it comes to the measurement of the amount of information of the memory structure of a module, the well-known Shannon formula (1) (1948) can be applied, but with the condition that the value coming out of this formula should be calculated in absolute values:

$$H = -\sum_{1}^{n} p_i \log_2 p_i \qquad (1)$$

Moreover, the other condition that has to do with the fulfillment or not of the value of the necessary function should be accepted. When the fulfillment value oscillates from 0 to 0.5, then the information amount of the memory module structure or the information amount of the module, which is to be marked with the letter 'i' will be calculated using the formula (2):

$$i = H \qquad (2)$$

and when this value is higher than 0.5 then the amount of information of the module is calculated through the formula (3):

$$i = 1 + (1-H) \qquad (3)$$

Let's analyze the case when the *lactase* enzyme which breaks down the milk sugar is found in 99 % of the individuals, that is the Swedish and in 1% of Asian-Americans individuals living in the USA. The question likely to be raised is: What is the amount of information in the Swedish population compared with the Asian-Americans? The answer depends on the application of formula (2) and (3); the application of the formula is valid only for a certain period and certain place. In the case of the Swedish population, which has consumed and is still consuming the milk and its by-products, the 0.99 frequency of individuals which have the *lactase* serves at the same time as the probability of the fulfillment of the necessary function. In this case we apply the formula (3), which leads to the idea that the amount of information of this module in the Swedish population is i = 1.92. On the other hand, in the case of the Asian-Americans, who now are milk consumers, the probability of the fulfillment of the function is quite low, only at 0.01, whereas the amount of information is i = H = 0.01.

But what is the amount of information in populations that do not consume milk? Is it safe to say that the amount of information in such a population is equal to 0? The answer is negative. In our opinion, in such case there cannot or never should be calculated the amount of information because there is no need or request to be taken care of it in the living organism. In this fact alone we should state that without the need of the task (0), there is no need of any function in the living systems; consequently one cannot speak of a value of information or its corresponding meaning. Even the reproduction modules have their one value of information. It is thought that the survival in certain bacteria is 0.00001%, while in mammals this figure stands about 20%, and within the human population it is fixed about 80% (in developed countries). Accepting the idea that the average percentage of the offspring in a given species is the fulfillment probability of the reproduction function module, then the amount of information of the reproduction module in mammals would be i = H = 0.72, while in the abovementioned human population it would be 1 = 1 + (1-H) = 1.72.

3. Definition of gene

The examples which helped to illustrate the module concept make us believe that the gene cannot be the only structure that carries information and which perform a function in the living systems.

The gene is a fragment of the nucleic acids, which in a certain type of module, called the genetic module, plays the role of a memory structure. It will be considered as a definition of gene that will also lead us to the thinking that the gene as a memory structure (S), taking the meaning only when its product or effector is an interpretation system (Sharov 2009); it would be able of solving an adaptive problem (O). Stated otherwise, the gene would be an information model (S) where after which the effector is formed is compared with a key (I), carrying out the necessary function, opening the respective lock (O).

The definition of gene, as given above, frees biology from the dogma that information which is used by the living systems is only the one of nucleic acids. Naturally, the biggest numbers of modules do not have as memory structures the fragments of nucleic acids.

Morever, from the modular concept we are able to understand that each function of the living systems is performed based in certain information under the conditions when a need or request to be satisfied arises in such living systems.

The information, according to which a function of living systems is fulfilled is meaningful and is totally different from that information, a concept which has been borrowed from physical or technical fields. The information that the memory module structures carry out, in which are included the genes as well, gains in value and meaning, it is calculated on the basis of fulfillment probability of the function.

The modular concept of the gene being in accordance with concept P (Moss 2008) and the ultimate causes of biological phenomena (Mayer 1961) make possible the difference with concept D and the proximate causes. The difference in concept between P (ultimate causes) and D (proximate causes) is shown in our modular concept of gene, through the difference between genetic modules and the non-genetic ones.

At the core of this difference is the idea that gene as a memory structure has extracted the information back in the past from the environment while the modules, in which the memory structure are not fragments of nucleic acids, living systems will yield just in the moment when they need to conserve their identity they have to carry out a definite function.

In this way, the modular concept of gene allows us to overcome several conceptual limitations. Above all with the presentation of the living systems through the modules networks it is possible to think about certain biology, probably not focused upon the genes, but based upon the information.

References

Bajrami Z. (2008), Essay on a Model of the Living Systems. Albanian Journal of Natural and Technical Sciences, Vol.14(23) 153-161p

Cosmides L. and Tooby J. (1992), Cognitive adaptation for social exchange.In: J.Barkow, L. Cosmides and J.Tooby (Eds).The adapted Mind:Evolutionary Psychology and generation of culture. New York, Oxford University Press

Dawkins R. (1982), The Extended Phenotype. Oxford University Press.

Emmeche C. (2002), The chicken and the Orphean egg: On the function of Meaning and the Meaning of function. Sign Systems Studies, 30 (1) 15-32p

Falk R. (2000), The gene: a concept in tension. In Peter Beurton, Rafael Falk and Hans – Jorg Rheinberger(eds) The concept of the Gene in Development and Evolution. Historical and Epistemological Perspectives. Cambridge. Cambridge University Press, 317- 348

Hoffmeyrer J. (2005), How can it be the case that the world contain human beings? The two cultures: Re-considering the decision between the sciences and humanities, Sydney. 21-22 July, 2005

Jonker C. M. and Treur J. (2006), Agent-Oriented Modeling of Dynamics of Biological Organisms.http//www.few.vu.nl-wai/Papers/APIO6bio.pdf

Maynard–Smith J. and Szathmary E. (1995), The Major Transition in Evolution. W.F. Freeman. Oxford. New York. Heidelberg.

Mayr E. (1961), Cause and effect in Biology. Sciences. 134, 1501-1506 pp.

Moss L. (2008), The Meaning of the Gene and future of the genotype. Genetics, Society and Policy, 4,38-57p

Noble D. (2008), Genes and Causation. Philosophical Transaction. Series A Mathematical, Physical and Engineering Sciences 366 (1878), 3001 – 3015 pp.

Peirce C. S. (1955) Philosophical Writing of Peirce. Ed. J.Buchler.New York:Dover Publications

Shannon C. (1948), A Mathematical Theory of Communication. Bell Systems Technical Journal 27: 279 – 423 pp.

Sharov A. A. (2009), Role of Utility and Interference in the Evolution of Functional Information. Biosemiotics, 2. 101-115 pp.

Uexkull J. (1982), The theory of Meaning. Semiotics. 42: 25 – 82 pp.