Eighth International Conference in Code Biology

Olomouc (Czech Republic) 7-11 June 2022

ABSTRACTS

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

Revisiting of "The Multiple Codes of Nucleotide Sequences"

Alexander Bolshoy
Department of Evolutionary and Environmental Biology
University of Haifa, Israel
bolshoy@evo.haifa.ac.il

In 1989 E. N. Trifonov published his highly-cited article "The Multiple Codes of NucleotideSequences". In this pioneer paper Trifonov introduces his vision of the term "sequence code": "Code is a sequence pattern instructive for one or another specific molecular (multimolecular) interaction or process". As a matter of fact, these "Trifonov's codes", discussed in several publications of Trifonov and his coworkers, are of very various nature. In some cases, authors deal with instructions (chromatin code, translation framing code, protein folding code), in other cases Trifonov and others are talking about polysemanticity of a string (overlapping codes). In some publications we deal with rather a method for detecting codes (contrast words), in other cases we seem to be closed to what semiotically a code really is: a pattern with a given syntax coding for a given content. Here are the codes presented in Trifonov's paper: Translation Framing Code; Chromatin Code; Shape Code; Loop Code. He also discusses "overlapping and degeneracy of the codes" and hidden meaning of tandemly repeated sequences. We would try to retell contents of this influential paper following two major criteria: a) we would try taking into account that our reader is not an expert in Molecular Biology and she/he is glad to get our 30-years-after linguistic/semiotic retelling the tale of sequence code multiplicity; b) research of genetic sequence has never stopped and we would like to inform our reader regarding modern widespread opinions in the field. One of the aims of this talk is to show how a universalistic view of "genetic codes", often introduced into molecular biology through physics, has been crumbling under the onslaught of evidence from various branches of molecular biology over the past thirty years of molecular biology research. We will try to draw conclusions from this collapse of universalistic concepts in molecular biology. Further we would like to discuss what makes a "point-of-view" paper inspiring and widely cited because in spite of bringing wrong or non-convincing examples of DNA codes in that Trifonov's famous paper, it has become soundly influential.

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

Genome-wide identification of high-affinity nitrate transporter 2 (*NRT2*) genes in tomato

Aydın Akbudak[1,*], Ertugrul Filiz[2], Durmuş Çetin[1]
[1]Akdeniz University, Department of Agricultural Biotechnology, Antalya, Turkey
[2]Duzce University, Department of Crop and Animal Production, Cilimli, Duzce, Turkey
*Corresponding author:
akbudak@akdeniz.edu.tr

The present study sought to identify the *NRT2* gene family members within the tomato genome, and then to characterize them in detail by means of bioinformatics, physiological and expression analyses. Four novel *NRT2* genes were identified in the tomato genome, all of which contained the same domain belonging to the major facilitator superfamily (PF07690). The co-expression network of the *SlNRT2* genes revealed that they were co-expressed with several other genes in a number of different molecular pathways, including the transport, photosynthesis, fatty acid metabolism and amino acid catabolism pathways. Several phosphorylation sites were predicted in the NRT2 proteins. The *SlNRT2* genes interact with many other genes that perform various functions in many crucial pathways within the tomato genome. The sequence variations observed at the gene and protein levels indicate the dynamic regulation of the *SlNRT2* gene family members in relation to cell metabolism, particularly with regard to the nitrogen assimilation pathway. The responses of the *SlNRT2* genes to drought and salinity stresses were found to be diverse, and they were neither stress- specific nor tissue-specific. The findings of this study should provide a useful scientific basis for future studies concerning the roles of the *NRT2* gene family in plants.
Key words: NRT2, gene family, expression profiles, tomato, abiotic stress

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

Paradigm for the Living
Evolutionary/Developmental Principles of Process and Inter-action

Anna Aragno
National Association for the Advancement of Psychoanalysis
140 West End Avenue, New York, NY, 10023, USA
annaragno@earthlink.net

Let us examine not how observation, facts and data, are built up into general systems of physical explanation, but how these systems are built into our observations, and our appreciation of facts and data. N.R Hanson, 1958. 3

What is feasible is the understanding of general forms. B. Russell 1918,109

This intermezzo presentation lays the foundations for a shift: from a fundamentally psychoanalytic-based body of knowledge to a wider evolutionary/developmental approach to the natural history of *signification*. Applying this new poly-disciplinary perspective, and to affect this radical conceptual shift, I have had to vastly expand my frame of reference superimposing new areas of research and information on all previous knowledge; in essence, create a synthesis. The Coronavirus lockdown of 2020 provided an opportunity to think and tackle new areas of study. It made clear that to go forward I had to go backward in time. So I armed myself with recent literature in cognitive and evolutionary-neuroscience, comparative developmental psychology, paleo-anthropology and psycho-history. Going forward I will apply this new body of knowledge, superimposed on a psychoanalytic background, to Code Biology's basic tenets and founding principles within the broad three-phase macro-evolutionary design of Organic, Neural, and Cultural forms of semiosis. I ask what a psychoanalytically informed approach may adduce from contemporary research on the evolution of the human brain/mind in terms of uncovering the generative sources and natural development of semiosis. Via conceptual reconstruction and inference, this poses the question; how did our singular species begin using 're-presentation,' devising carriers of signification as instruments of mind. To approach this and related topics methodically I needed to lay out a set core of principles for a new paradigm, which I herewith present. This, and presentations going forward, form an interconnected series building towards a summarizing synthesis. Here, I lay out a detailed list of core principles guiding a developmental/evolutionary approach applicable to select topics within an overarching Barbierian framework of three macro-evolutionary shifts.

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

Traumas, Triggers, and Codes: Applications of Code Biology in the Theory of Addiction

Candice Shelby
University of Colorado, Denver, CO 80217-3364
Candice.Shelby@ucdenver.edu

This paper examines the function of several levels of codes in the phenomenon of addiction. In my 2016 book I argued that the standard disease model of addiction (I use the term 'addiction' as an abbreviation, since it is well documented that there is no single situation captured by that term), falls far short of accounting for the phenomenon. In general, the scientific community (and certainly the funding agencies in the U.S.) maintain that addiction is best characterized as a disregulation of dopamine pathways in the nucleus accumbens/ventral tegmentum areas of the brain, such that sustained high levels of dopamine result in the loss of dopamine receptors, which then requires ingestion of greater amounts dopamine-producing or dopamine-preserving substances or activities (gambling was the first of these to be identified) in order to achieve the same positive experience, or even to minimize negative ongoing experience. This simple mechanistic account both fails to account for the phenomenon and completely ignores the crucial role played by biological codes at every level. Biological codes are involved in addiction from the start, since the genetic code accounts for approximately 50% of the likelihood that an individual will experience addiction. Additionally, the scientific account of addiction fails because it only explains dependence. Addiction is much more than developing dependence. Many people who have painful surgeries, for instance, can become dependent on substances, in particular on opioids, such that they experience discomfort when the drug is withdrawn, as their depleted opioid-receptors adjust. But most of these patients will simply accept their increased discomfort as an expected features of surgery and lessening medication. Individuals who experience addiction, however, on their own accounts, remain sensitive to the meanings of objects and situations around them in a particular way. They can be "triggered" by a smell, a particular music, a landscape, or some other arbitrary thing that is coded in their psychological lives to stimulate use of the substance or performance of the activity to which we often say that they are addicted. And this happens regardless of whether or not they have been involved with that substance or activity for months, or years, or decades. So not only is this not dependence; it is also *not*, I will argue, explained by a simple mechanism, as some other "triggers" can operate, such as certain triggers in a trauma response, which are clearly representative of the trauma-causing phenomenon, and so set into action similar physiological reactions to those caused by the original trauma. Psychologically *coded* "triggers" in the sense that I wish to draw attention to are distinct from those; the ones alluded to here depend on a transmission of an arbitrary signal that requires interpretation to achieve its function. Finally, and perhaps most importantly, social codes play a powerful role in stimulating or repressing what we might call addictive behaviors, as well as in determining which behaviors we characterize as addictive. Philosophers, for example, will often revert to smoking cigarettes at a conference (there remains a fateful event at the American Philosophical Association's conferences known as "the smoker") even those who have no dependence on nicotine, nor do they usually ever even think about smoking. More urgently, as adduced by the 22% increase in overdose deaths in the U.S. from 2020 to 2021, to a total of more than 100,000 deaths in a single year, social codes are indicated significantly in both the account of what we agree to label as addiction, and in the circumstances that may lead to an untethered desire to self-medicate. Here I refer to social coding about who among us count as relevant or valued parts of society, and perhaps in what is considered reasonable self-soothing relative to social conditions. At all levels, then, from the organic to the social, biological codes provide powerful tools for explaining the true nature of addiction.

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

Robustness against point-mutations of genetic code extensions

Elena Fimmel and Lutz Strüngmann
(joint work with Markus Gumbel and Martin Starman)
Institute for Mathematical Biology, Faculty of Computer Science,
Mannheim University of Applied Sciences, 68163 Mannheim, Germany
e.fimmel@hs-mannheim.de
l.struengmann@hs-mannheim.de

Many theories of the evolution of the genetic code assume that the genetic code has always evolved in the direction of increasing the supply of amino acids to be encoded. In order to reduce the risk of the formation of a non-functional protein due to point mutations, nature is said to have built in control mechanisms. Using graph theory Blazej et al. investigated if this robustness is optimal in the sense that a different codon-amino acid assignment would not generate a code that is even more robust. At present, efforts to expand the genetic code are very relevant in biotechnological applications, for example, for the synthesis of new drugs. In this talk we generalize the approach proposed by Blazej et al. and will explore hypothetical extensions of the standard genetic code with respect to their optimal robustness in two ways: (1) We keep the usual genetic alphabet but move from codons to longer words, such as tetranucleotides. This increases the supply of coding words and thus makes it possible to encode non1canonical amino acids. (2) We expand the genetic alphabet by introducing non-canonical base pairs. In addition, the approach from Blazej et al. is extended by incorporating the weights of single point mutations into the model. The weights can be interpreted as probabilities (appropriately normalized) or degrees of severity of a single point mutation. In particular, this new approach allows us to take a closer look at the wobble effects in the translation of codons into amino acids. According to the results from Blazej et al. the standard genetic code is not optimal in terms of its robustness to point mutations if the weights of single point mutations are not taken into account. After incorporation into the model weights that mimic the wobble effect, the results of the present work show that it is much more robust, almost optimal in that respect. The results were published in [1] and supported by computational results in [2].

[1] Fimmel E., Gumbel, M., Starman, M., and Strüngmann L.: Robustness against point mutations of genetic code extensions under consideration of wobble-like effects, Biosystems (2019).
[2] Fimmel E., Gumbel, M., Starman, M., and Strüngmann L.: Computational Analysis of Genetic Code Variations Optimized for the Robustness against Point Mutations with Wobble-like Effects, Live (2021).

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

"Double-agents" in oncogenesis – when genes can play on both sides

Hana Owsianková
Department of General Linguistics, Palacký University
Olomouc, Czech Republic
hana.owsiankova@upol.cz

As defined by molecular genetics, a gene is a DNA segment that encodes a functional product where in the functional product is meant RNA or its translation product, a protein (a polypeptide). The function of these products is directly related to their structure. In this contribution I would like to focus on the structure of the genes which encode products with ability to switch their function from the default one to the right opposite, based on cellular context. In particular, my aim is to discuss this structure-function peculiarity observed among human genes whose changes in expression usually cause cancer. Simply put, cancer is a designation for a large group of diseases manifesting an abnormal cell growth and the potential of spreading to other parts of body – not only environmental but also genetic factors play crucial role in cancer development. In the cells of multicellular organisms we can find both oncogenes and tumor-suppressor genes – oncogenes can cause normal cell to grow out of control and become cancer, tumor-suppressor genes protect them from degenerating into cancer cells. In carcinogenesis, they seem to be contradictory gene groups, yet some genes have dual roles, i.e. display both oncogenic and tumor-suppressor functions under different cellular context (see Soussi – Wiman, 2015; Yang et al., 2007; Yip et al., 2010). Because we are discussing topic concerning genetic code and encoding genetic information, the structure of above described genes will be examined by quantitative methods that primarily fall within the field of linguistics– a field which, besides other things, deals with codes of all types. These methods, e.g. Damerau-Levensthein distance, Bag-of-words model, Manzerath-Altmann law, have been introduced to genetics more recently (see Baixeries et al., 2013; Bolshoy, 2003; Lovato, 2015; Sutton et al, 2014).

References

Baixeries, J. – Hernández-Fernández, A. – Forns, N. – Ferrer-i-Cancho, R. (2013). The Parameters of Menzerath-Altmann Law in Genomes. *Journal of Quantitative Linguistics* 20(2), pp. 94–104.

Bolshoy, A. (2003). DNA sequence analysis linguistic tools: contrast vocabularies, compositional spectra and linguistic complexity. *Applied bioinformatics* 2, pp. 103–112.

Lovato, P. (2015). *Bag of words approaches for Bioinformatic*, Ph.D. thesis, Dept. of Computer Science, University of Verona, series TD-03-15.

Soussi, T. – Wiman, K. G. (2015). TP53: an oncogene in disguise. *Cell Death Differ.* 22, pp. 1239–1249.

Sutton, L. A. et al. (2014). An Entity Evolving into a Community: Defining the Common Ancestor and Evolutionary Trajectory of Chronic Lymphocytic Leukemia Stereotyped Subset #4. *Molecular Medicine* 20(1), pp. 720–728.

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

Objective, Immediate, Initial, and Dynamical Interpretants: Which one of Peirce's semiotics is compatible with code biology?

Ivan Fomin
Palacky University, Olomouc, Czech Republic
fomin.i@gmail.com

Semiotic theory of Charles Peirce tends to be associated with "nonmechanistic" approach in biosemiotics that implies that any meaning is the result of interpretation. This perspective is opposed to the "mechanistic" approach of Code Biology which suggests that meaning is a result of coding. The contrast between the two approaches is important, however the fact that Peirce's semiotics has to be exclusively associated with the "nonmechanistic" one is not obvious. The elements of Peirce's theory of signs are scattered across the corpus of his writings, which is far from being terminologically and conceptually consistent, so one can hardly speak of just one Peircean semiotics, but rather about different possible reconstructions of different versions of semiotics that Peirce outlined in his various texts. The question is if there can be versions of Peirce-inspired semiotics that do allow for meaning without abduction and interpretation. If such reconstructions of Peircean semiotics are possible, they can emerge from an analysis of Peirce's diverse attempts to define the notion of interpretant and distinguish between different kinds of interpretants. Even though the very form of the term "interpretant" suggests its connection to interpretation, one can find that sometimes Peirce uses this term to refer to just "anything that the sign, as such, effects" (ILS 285) without suggesting that the effect is necessarily characterized as abductive, mental, or interpretive. Moreover, in several versions of the typology of interpretants Peirce talks of different kinds of effects that may be considered as interpretants, including, for example, immediate interpretants, defined as mere "qualities of impression that a sign is fit to produce" (CP 8.315), or dynamical interpretants, defined as "actual events" brought about by signs (ILS 285).

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

How the cell manufactures itself

Jan-Hendrik S. Hofmeyr
Department of Biochemistry, University of Stellenbosch
Stellenbosch 7600, South Africa.
jhsh@sun.ac.za

Being fragile, yet persistent, living longer than the functional lifetimes of one's components: these hallmarks of life are made possible by the remarkable ability of living organisms to manufacture and individuate themselves autonomously as wholes, an ability that is arguably the most fundamental property that distinguishes living from non-living systems. In order to grow, reproduce, metabolise, self-maintain, adapt and evolve, living organisms must first and foremost be able to self-manufacture. To understand how the cell accomplishes this feat, I divide cellular processes into three classes of efficient causes that produce each other, so making the cell closed to efficient causation, the hallmark of an organism. These classes are the *enzyme catalysts* of covalent metabolic chemistry, the *intracellular milieu* that drives the supramolecular processes of chaperone-assisted folding and self-assembly of polypeptides and nucleic acids into functional catalysts and transporters, and the *membrane transporters* that maintain the intracellular milieu, in particular its electrolyte composition. Each class of efficient cause can be modelled as a relational diagram in the form of a mapping in graph-theoretic form, and a minimal model of a self-manufacturing system that is closed to efficient causation can be constructed from these three mappings using the formalism of relational biology [1]. This Fabrication-Assembly or *FA*-system serves as an alternative to Robert Rosen's replicative Metabolism-Repair or *MR*-system, which has been notoriously problematic to realise in terms of real biochemical processes. A key feature of the model is the explicit incorporation of formal cause, which partly arrests the infinite regress that characterises all relational models of the cell. The *FA*-system is extended into a detailed relational model of the self-manufacturing cell, a model that has a clear biochemical realisation.

[1] Hofmeyr, J.-H.S. (2021) A biochemically-realisable relational model of the self-manufacturing cell. BioSystems 207, 104463.

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

Dreams as manufactured mental processes

João Carlos Major
International Academy of Analytical Psychology
Braga, Portugal
jcmajor@mail.telepac.pt

The author argues that we can study the 'unconscious mind' or 'emotional brain' in the scientific framework of 'Code Biology' (understood as *the study of all semiotic processes of life with the standard methods of science*). The most prominent aspect of the unconscious mind is dreams. Dreams are 'manufactured mental processes,' where the keyword is 'manufactured' because this word means 'produced according to rules,' i.e., 'according to codes.' Dreams are 'autopoietic' mental processes manufactured according to 'dream codes.' In reality, dreams are produced by what also causes the awake brain: the unconscious or emotional mind. The awake brain or conscious mind is a fine-tuning of the unconscious mind; in fact, only a tiny portion of the unconscious brain's myriad processes. And because the unconscious brain is operating at a much more emotional level than when we are awake, our emotional brain makes much more connections regarding our feelings and experiences than our conscious self. So, dreams may be ways of confronting emotional dramas, breaking up, and reassembling the events according to pre-existing adaptative emotional rules– the 'dream codes.' This, in turn, means that the awake brain is working upon emotional or unconscious principles, and we have an evolutionary advantage if we are aware of and find out what these 'brain codes' (or archetypes) are, and this is a thread that is worth thinking about.

Bibliography

Major, J.C. (2022). Archetypes and Code Biology. *BioSystems*, 208: 104501. https://doi.org/10.1016/j.biosystems.2021.104501

Gauchat, A. et al. (2020). Disturbing Dreams and Psychosocial Maladjustment in Children: A Prospective Study of the Moderating Role of Early Negative Emotionality. *Frontiers Neurology*. https://doi.org/10.3389/fneur.2020.00762

Barbieri, M. (2019). A general model on the origin of biological codes. *BioSystems*, 181: 11-19. https://doi.org/10.1016/j.biosystems.2019.04.010

Samson-Daoust, E.,et al. (2019). Predicting the affective tone of everyday dreams: A prospective study of state and trait variables. *Scientific Reports*. DOI: 10.1038/s41598-019-50859-w

Demacheva, I. &Zadra, A. (2019). Dream content and its relationship to trait anxiety. *International Journal of Dream Research*, 12(2): 1-7.

Stickgold, R. & Walker, M. (2007). Sleep-Dependent Memory Consolidation and Reconsolidation (2007). *Sleep Medicine*, Jun8(4): 331–343.

Franklin, M. &Zyphur, M. (2005). The Role of Dreams in the Evolution of the Human Mind. *Evolutionary Psychology,*3: 59-78. https://doi.org/10.1177/147470490500300106

Pesant, N. &Zadra, A. (2004). Working with dreams in therapy: What do we know and what should we do? *Clinical Psychology Review,* 24: 489-512

Maturana, H. (1987). The biological foundation of self-consciousness and the physical domain of existence. In E.R. Caianillo (Ed.).*Physics of process*. Singapore: World Scientific, 324-379.

Jung, C. G. (1969). The Structure and Dynamics of the Psyche. *Collected Works*, Vol. VIII. Princeton, NJ: Princeton University Press.

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

Codegenesis
Harnessing Higher-order Work Pathways via Annealing Networks

Joshua Augustus Bacigalupi
Dactyl Foundation for the Arts and Humanities
64 Grand Street, New York, NY 10013 USA
bacigalupiworks@gmail.com

This presentation explores how conventions, such as codes, emerge, stabilize and complexify themselves in living systems. The philosophical frame and motivations for this investigation will be illustrated by a series of general questions about the intrinsically generative nature of life, from which an initial intuition will be proposed. This intuition is the basis upon which a more rigorous theoretical model is proposed. This theory of how codes generate and sustain themselves will then serve as a heuristic lens through which modern mathematical and empirical investigations will be elucidated as possible practical implementations of the proposed theory of *codegenesis*.

A Question and an Intuition

Starting with questions inspired by an architectural design process: how are numerous constraints and parameters that impact the eventual design resolved? These initial constraints are often contradictory, but, through a deliberate process, these often dissonant considerations are concretized into a complex, coherent and singular code able to implement complex dynamics via extant cultural institutions. But how is such generative creativity able to create a code that is then capable of realizing adaptive outcomes well beyond the code itself? Inspirations from quantum mechanics and physical wave mechanics provided an intuition: *superposed interference patterns*.

Codegenesis in Theory

Building on thermodynamic and information theories, the above questions and initial intuitions undergird a rigorous theory. It proposes a population of diversely tuned nodes, analogous to neurons, that rewire themselves via sparse higher dimensional interference within the superposed noise of their shared common medium; this global interference behavior molds and is molded by each nodes' local behavior, *simultaneously*. And through a self-regulatory annealing habit, ever more complex and semiotically relevant codes can emerge. The theory proposes that the emergent metastable dynamics of this model may be tested for codegenesis via empirical parameters such as adaptive work pathways and accrued network complexity.

Codegenesis in Practice

This theory is used as a heuristic lens to find existing models of empirical behaviors in living systems able to exemplify and potentially implement codegenesis in future experiments. Such candidate phenomena include, but are not limited to, anticipatory amoeboid behavior explained via an extension of the Kuramoto model; reaction-diffusion and bioelectric codes used to explain morphogenesis; and self-regulatory extensions of the histone code used to explain cell differentiation.

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

Molecular Recognition Theory applied in biological experimental model: Reversal of the hepatoprotective effect of D-met-enkephalin by a complementary peptide

Katarina Matković
Institute for Medical Research and Occupational Health, Zagreb, Croatia
kmatkovic@imi.hr

Most regulatory processes in biological systems depend on protein-protein or protein-peptide interactions. Our knowledge of such interactions is relatively limited, and experimental methods for studying them are complex, slow and expensive. Therefore, we use heuristic methods to determine the bioactive regions in proteins and model interactions of different receptors/acceptors with their complementary (antisense) peptides [1]. One such method, Molecular Recognition Theory (MRT), is based on the Code Biology principles [2,3] and it has been practically applied in the experimental modeling of D-met-enkephalin hepatoprotection. Protective potential of D-met-enkephalin (D-YGGFM) was evaluated using an experimental model of paracetamol induced hepatotoxicity in male CBA mice and observed by using three criteria: plasma activities of ALT and AST enzymes, liver necrosis score and number of spontaneously dead animals [4]. The complementary peptide of D-met-enkephalin (IPPKY) was constructed using the MRT based translation algorithm in 3'→5' direction. The basic properties of the antisense peptide IPPKY and its interaction with D-met-enkephalin were studied using circular dichroism spectroscopy and fluorescence spectroscopy. In biological experimental model complementary (antisense) peptide IPPKY completely reversed the protective effects of D-met-enkephalin. It has been proven that using complementary sequences it is possible to construct a ligand-acceptor peptide system in a simple and efficient way.

References

1. Štambuk N, Konjevoda P, Pavan J. Int. J. Mol. Sci. 2021; 22:9106.
2. Štambuk N, Konjevoda P, Boban-Blagaić A, Pokrić B. Theor. Biosci. 2005; 123:265.
3. Štambuk N, Konjevoda P, Turčić P, et al. BioSystems 2018; 164:199.
4. Turčić P, Štambuk N, Konjevoda P, et al. Medicinal Chemistry 2015; 11:286.

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

The Lower Threshold in Code Biology

Ľudmila Lacková
Palacký University, Department of General Linguistics
Olomouc, Czech Republic
ludmila.lac@gmail.com

Code Biology concentrates on the notion of code: living organisms are defined as code-users or code-makers. In this presentation I introduce the 'lower coding threshold' as the boundary between coding and purely physical or chemical phenomena. The notion is inspired by the "semiotic" threshold, based on the influential book by Umberto Eco *A theory of semiotics* where he distinguishes two basic semiotic thresholds: lower and upper semiotic threshold. For Eco, the lower semiotic threshold is clearly defined, and it is situated at some point between signals and signs. Signals should not, according to Eco, make part of the semiotic inquiry. Signals should be excluded from the semiotic threshold zone because they are not dependent on conventionality and arbitrariness, yet are only definable in terms of dyadic stimulus-answer principle. For Eco, all biological phenomena are signals. As a consequence, no biological phenomena should be part of the semiotic project: neither at the genetic, epigenetic, cellular, nor neural level. In this presentation, I will confront the notion of signal and the notion of code. I focus on establishing the lower threshold at protein biosynthesis, and I propose basing the semiotic understanding of the lowest life forms on the following criteria: arbitrariness, representation, repetition, historicity and self-replication. I also offer that this definition of the lower threshold need not include the notion of interpretation, in the hope that this newly specified principle of the lower threshold can be accepted as a way forward in the conversation between Code Biology and Biosemiotics.

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

The Semantic Theory of Evolution: a case study of the revolution in science

Lukáš Zámečník
Palacký University in Olomouc
lukas.zamencnik@upol.cz

In the first half of the 1980s, Marcello Barbieri introduced his ribotype theory on the origin of life (Barbieri 1981) and subsequently supplemented it with a semantic theory of evolution (Barbieri 1985). Together, the two represented a fundamental change in molecular biology by diverging from a simple mechanistic conception of the code and introducing extended mechanisms that incorporate organic meaning – moving from a genotype-phenotype duality to a genotype-phenotype-ribotype trinity. Over time, Barbieri's theories have found several coherent forms (e.g. Barbieri 2003) in the last variant in contemporary code biology (Barbieri 2015, 2019). Karl Popper (Popper 1981, 1983) pointed out (along with René Thom, Jean Brachet, and Carl Woese) the revolutionary nature of Barbieri's semantic view of evolution (and especially ribotype theory on the origin of life). Barbieri built the semantic theory of evolution following the contemporary model-based view of theories (e.g. Suppes 1972, Suppe 1977, van Fraassen 1980) and precisely in the spirit of Popper's methodological maxims (Popper 1959, 1962). He built the theory as a bold conjecture that offers guidelines to test its relevance – looking for ways to refute it. Barbieri unequivocally supports this Popperian inspiration (Barbieri 1981).

Objectives of the paper:

1) The paper aims to examine this scientific revolution in more detail from the perspective of Popper's philosophy of science. We will conceive of the birth of the semantic theory of evolution as a case study of the use of Popper's methodological maxims. We will compare this case study with others, especially physical ones, contained in Popper (Popper 1959, 1962).

2) Using personal correspondence and reviews (Popper 1981, 1983) related to Barbieri's book (Barbieri 1985), we will see how this case study Popper himself introduces, why he highlights the ribotype theory, and what reservations he brings to the semantic theory of evolution.

3) The paper also aims to evaluate what falsification tests the semantic theory of evolution or code biology has undergone. On that occasion, we note Barbieri's research papers and books (see Barbieri 2019, 2015, 2003), which show code biology as a corroborated theory.

Literature:

Barbieri, M. (1981) The ribotype theory on the origin of life. Journal of Theoretical Biology, 91, 545–601.
Barbieri, M. (1985) The semantic theory of evolution. London/New York: Harwood Academic Publishers.
Barbieri, M. (2003) The organic codes: an introduction to semantic biology. Cambridge: Cambridge University Press.
Barbieri, M. (2015) Code biology: a new science of life. Dordrecht: Springer.
Barbieri, M. (2019) Evolution of the genetic code: the ambiguity-reduction theory. BioSystems, 185, 1–7.
Popper, K., R. (1959) Logic of scientific discovery. London: Hutchinson &Co.
Popper, K., R. (1962) Conjectures and refutations: the growth of scientific knowledge. London: Routledge.
Popper, K., R. (1981) Personal correspondence with M. Barbieri, I.
Popper, K., R. (1983)Personal correspondence with M. Barbieri, II.
Suppe, F. (1977) The Structure of Scientific Theories. Urbana: University of Illinois Press.
Suppes, P. (1972) Axiomatic Set Theory. New York: Dover Publications.
van Fraassen, B., C. (1980) The Scientific Image. Oxford: Oxford University Press.

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

Endless Forms of Endless Formation

Marcella Faria Almeida Prado
Dactyl Foundation for the Arts & Humanities
almeidapradomarcella@gmail.com

The present work takes the closing remark of "The Origin of Species" as its point of departure. In order to grasp the distinctive properties of matter organization in living forms as opposed to inanimate assemblies, we will focus on aspects of organic formation. Only biological objects come into being, they are morphed by the growth, development and evolution of simpler and smaller unities into functional wholes of higher complexity. Darwin ends his seminal book by wondering how "*whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved*". We believe that by choosing the term "endless", not surprisingly a polysemic word, he has provided the basis for the further development of his core idea into new (form)ulations. The pragmatic dimension (potential in action) is a paramount one in evolution, regardless of the substrate, be it acting upon an organism or an abstract formalization. To be actualized in context, a scientific concept, for example, must respond to a myriad of challenging inputs: the changing environment, the multiscale and hierarchical nature of its objects, the cumulative amount of experimental data to be integrated, the increasingly complex set of rules and conventions constraining its uses, the historical, and often-tacit, epistemic assumptions and interests that scientific practices. Here, we propose four alternative – even though highly complementary – meanings for the Darwinian idea of "endlessness", our motivation is to give that idea a wider and deeper explanatory power by incorporating theoretical inputs that came later. The four categories of infinitude we see as singularities of the living are: 1) endless, as in multiple and unpredictable forms; 2) endless, as in unfinished and undetermined forms; 3) endless, as forms which's end is their own formation; 4) endless, as forms which are able to reproduce themselves. In conclusion, we turn to mechanism, describing the processes at work at each scenario, which are respectively: 1) natural variation, natural selection, natural conventions, all leading to evolution; 2) ambiguity, degeneracy, emergency, epigenetics, all leading to meaning; 3) homeostasis, autopoiesis, self-manufacturing, codepoiesis, poiesis, all leading to stasis; 4) genetic coding, genetic copying, mutation, information, all leading to identity. Of notice, the sources of conceptualization we have used as background include a wide range of theoretical perspectives. We draw mainly on Code Biology, but also incorporate the literature produced by other disciplines, i.e., Relational Biology, Experimental Medicine, Philosophy of Sciences, Biosemiotics, Cognitive Sciences, Complexity Theory, Conceptual Metaphor Theory.

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

Introduction to Code Biology

Marcello Barbieri
Dipartimento di Morfologia ed Embriologia
Via Fossato di Mortara 64a, 44121 Ferrara, Italy
brr@unife.it

Various experimental discoveries have shown that many organic codes exist in living systems and this means that they came into being throughout the history of life. The genetic code appeared in a population of primitive systems called *common ancestor*, and the signal transduction codes were associated with the split of the common ancestor into Archaea, Bacteria and Eukarya. After the genetic code and the signal trasduction codes, the prokaryotes did not evolve new codes whereas the eukaryotes continued to explore the coding space and gave origin to splicing codes, histone code, tubulin code, compartment codes and many others. The prokaryotes, on the other hand, did not increase the complexity of their cells whereas the eukaryotes gave origin to increasingly complex organisms and this suggests that there is a close link between codes and complexity. Another important implication comes from the fact that there are *two* distinct molecular mechanisms at the basis of life, the *copying* of the genes and the *coding* of proteins. The first leads to *natural selection* and the second to *natural conventions*. The existence of copying and coding at the molecular level, in other words, means that there are two distinct mechanisms of evolution: *evolution by natural selection*, based on copying, and *evolution by natural conventions*, based on coding. This is the major concept of Code Biology: the existence of many organic codes in life means that there are two distinct molecular mechanisms in life and two distinct mechanisms of evolution.

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

The effect of combined application of tazemetostat and cisplatin is cell-type specific

Marko Tomljanović
Laboratory for Epigenomics, Ruđer Bošković Institute, Zagreb, Croatia
mtomljan@irb.hr

Enhancer of Zeste Homolog 2 (EZH2) is a histone lysine methyltransferase, indispensable for establishment of a transcriptionally repressive mark – trimethylation of lysine 27 in the histone H3 (H3K27me3). This mark plays an essential role in a proper cellular differentiation and maintenance of stem cells pluripotency. In addition to these physiological roles, EZH2 plays and important role in oncogenesis. Elevated levels of EZH2 have been shown in several types of cancer, which is associated with repression of tumor suppressor genes. Currently, EZH2 is considered to be an oncogene and it is becoming an attractive target for pharmacological inhibition. Currently, there is one available EZH2 inhibitor, tazemetostat (Tazverik®), which is approved for the treatment of follicular lymphoma and epithelioid sarcoma. There are a few other EZH2 inhibitors which are being explored in various phases of clinical trials. Several in vitro studies, primarily based on lung and ovarian cell culture models, have shown the benefit of EZH2 inhibition combined with conventional antitumor drugs, such as cisplatin.
In our laboratory we are exploring the effect of tazemetostat in combination with cisplatin on, so far, unexplored cellular models, using the cell lines originating from head and neck cancer and colon cancer. Preliminary results have revealed several interesting effects that are not in agreement with the data published so far. The modulation of EZH2 will be discussed considering the histone code [1,2].

References
[1] Kühn S, Hofmeyr, J-HC. Is the "histone code" an organic code? Biosemiotics 2014; 7:203.
[2] Barbieri M. Overview of the third special issue in code biology. BioSystems 2021; 210:104553.

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

The wobble-effect and its Influence on genetic code variations optimized for the robustness against point mutations – a computer analysis

Markus Gumbel
(joined work with Elena Fimmel, Martin Starman and Lutz Strüngmann)
Center for Algorithmic and Mathematical Methods in Medicine, Biology, and Biotechnology,
Mannheim University of Applied Sciences, 68163 Mannheim, Germany
m.gumbel@hs-mannheim.de

Codon to amino acid assignments of the standard genetic code might help to minimize problems caused by point mutations. The robustness of a code against point mutations can be described with the help of a so-called conductance measure [1] – a weighted graph-based method. This presentation analyzes the influence of the wobble effect on genetic code tables and seeks for optimal robustness using an evolutionary optimization algorithm [2, 3] which optimizes the weights of the conductance graph. We demonstrate that the robustness is least influenced by mutations in the third position—like with the wobble effect. The results clearly demonstrate that point mutations in the first, and even more importantly, in the second base of a codon, have a very large influence on the robustness of the genetic code. These results are put in context to single nucleotide variants (SNV) in coding sequences. The question is addressed which structure of a genetic code evolves from random code tables when the robustness is maximized. Our results illustrate that the evolving code tables are very close to the standard genetic code that the robustness against point mutations seems to be an important factor in the evolution of the standard genetic code.

[1] Błażej, P., Kowalski, D.R., Mackiewicz, D., Wnetrzak, M., Aloqalaa, D.A., Mackiewicz, P., 2018. The structure of the genetic code as an optimal graph clustering problem (preprint). bioRxiv. https://doi.org/10.1101/332478
[2] Fimmel, E., Gumbel, M., Starman, M., Strüngmann, L., 2021. Robustness against point mutations of genetic code extensions under consideration of wobble-like effects. Biosystems 208, 104485. https://doi.org/10.1016/j.biosystems.2021.104485
[3] E. Fimmel, M. Gumbel, M. Starman, L. Strüngmann: Computational Analysis of Genetic Code Variations Optimized for the Robustness against Point Mutations with Wobble-Like Effects. Life (2021), 11, 1338. https://doi.org/10.3390/life11121338

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

I-Circular Codes

Martin Starman and Lutz Strüngmann
Mannheim University of Applied Sciences,
Paul-Wittsack-Str. 10, 68163 Mannheim, Germany
m.starman@hs-mannheim.de

A message like mRNA, which consists of continuous characters without separators (like commas or white spaces), can easily be decoded incorrectly if read in the wrong reading frame. A construct to theoretically avoid these reading-frame-errors is the class of block codes. However, already the first hypothesis by Watson and Crick [3] that block codes are used as a tool for reading-frame-error prevention in coding sequences failed. The problem with their hypothesis is that the four identity codons AAA, CCC, GGG and UUU seem to play an important role in protein coding sequences. Even the class of circular codes, which was discovered in coding sequences by Arquès and Michel [1], cannot include an identity codon. Yet, by including the interpretation of the message into the reading-frame robustness, the extension of circular codes to include identity codons is theoretically possible. Therefore, we introduce the new class of I-circular codes. Other than circular codes, these codes allow frame shifts, but only if the decoded interpretation of the message is identical to the intended interpretation. In the following, we introduce the formal definition of I-circular codes as well as the maximal and the maximum size of I-circular codes using the standard genetic code table. These numbers are calculated using a new graph theoretical approach derived from the class of circular codes. Further, we will show that all 216 maximally self-complementary C 3 codes (see [2]) can be extended to longer I-circular codes. Based on these codes, we present the increased code coverage of the 216 newly created I-circular codes based on the human coding sequences in chromosome 1. In the final part of this project we use the polarity of amino acids as an interpretation table to construct I-circular codes. In an optimisation process, two maximum I-circular codes of length 30 are found.

[1] D. Arquès and C. Michel. A complementary circular code in the protein coding genes. Journal of Theoretical Biology, 182(1): 45–58, September 1996.
[2] E. Fimmel, S. Giannerini, D. Gonzalez, and L. Str¨ungmann. Circular codes, symmetries and transformations. Journal of Mathematical Biology, 70(7): 1623–1644, 2015.
[3] J. Watson and F. Crick. Molecular structure of nucleic acids. Nature, 171(4356): 737–738, 1953.

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

On the inconsistency of the stereochemical models for the origin of the genetic code

Massimo Di Giulio
The Ionian School, Genetic Code and tRNA Origin Laboratory,
Via Roma 19, 67030 Alfedena (L'Aquila), Italy
massimodigiulio1952@gmail.com

Here, I analyse some circumstances under which the stereochemical theory of the origin of the genetic code would seem to be unable to provide a natural model for the origin of the code. For example, the fact that this theory predicts an interaction between the codon (or anticodon) and the amino acid for the origin of the genetic code would appear to be unsatisfactory because the coding of amino acids in the genetic code would only represent an intermediate step towards the real coding represented instead by genes coding for proteins. In other words, from a good stereochemical theory one would expect instead that this interaction goes directly to define the final product, that is to say proteins (genes), performing the real task in the evolving biological system and not the intermediate one defined by most of stereochemical models. It is therefore seen that stereochemical models would not seem able to provide a natural response to the origin of the genetic code because they would not clearly and directly define the origin of genes, that is, the origin of the mRNA. I also consider other arguments which would seem to limit the likelihood of the stereochemical theory.

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

Outlines for theorizing codes and coding

Mikhail Ilyin
Center for Advanced Methods in Social Sciences and Humanities
Russian Academy of Sciences, Moscow, Russia
mikhaililyin48@gmail.com

The word *code* is a catchword and a slogan (from Gaelic *sluagh +gairm*, or battle cry) of code-biology. This English word and its equivalents or rather transliterations in other languages have many meanings. Various dictionaries typically give from half a dozen to a dozen basic meanings. Strictly speaking, it is not a single word but rather an assortment of homonymous words. They form a cluster united by family resemblance just like the words *game* and *play* – c.f. famous example by Ludwig Wittgenstein in his *"Die Philosophischen Untersuchungen"* (paragraph 66 ff). In the same way as the words *game* and *play* (*das Spiel*) have affinity (resemblance) due to a common functional load of exercise of vital human capacities the words *code* gain affinity thanks to involvement into processes of coding and decoding.

The paper resorts to ordinary language analysis or Sprachspiel. It addresses a set of questions: What is the role of codes in coding and decoding? Do they play the roles of an agency, a medium or media, a vehicle, an instrument or anything else? Analysis proves that different words *code* in different cases of word-usage refer to varieties of different aspects and properties of coding/decoding, as well as to their outcomes, prerequisites or other related entities. The paper also considers a variety of claims about codes and their role in coding/decoding or communication at large, e.g., the famous dictum by Marshall McLuhan – the medium is the message. This analysis makes it possible to single out alternative of models of a variety of entities habitually called *codes*. By and large interpretations of such entities (prerequisites, agents, instruments, vehicles, media, outcomes etc. of coding/decoding) prove their specific functional character and quality, but at the same time their interdependence and affinity.

The paper strives to construct an interconnected assemblage of alternative models in a form of a network of functional linkages. It places at the core a pure logical relational scheme (bare code) and surrounds it with structurally similar models of complementary aspects and constituents of coding/decoding in a kind of fractal-like replications. The core (bare code) jointly with its alternative extensions constitute bigger complex entities called extended codes. This pilot logonomic setup is far from being complete. Furthermore, it still remains chiefly a kind of methodological guide or blueprint for elaboration of a far more complex and authentic construction of a comprehensive prototype model of coding/decoding anchored in domains of communication from biological and lingual ones to political and economic.

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

The Molecular Recognition Theory Revisited: The Relational Model of the Standard Genetic Code

Nikola Štambuk and Paško Konjevoda
Ruđer Bošković Institute,
Bijenička cesta 54, 10000 Zagreb, Croatia
stambuk@irb.hr

The standard genetic code (SGC) describes an algorithm for the translation of DNA and RNA information into proteins. The natural property of this genetic algorithm is the complementarity principle, i.e., pairing of uracil (U) or thymine (T) with adenine (A), and cytosine (C) with guanine (G). Complementary amino acid coding is specified by four main rules [1]:

1. Direction of DNA and RNA sequence translation unambiguously defines differences in: the number of complementary amino acids pairs (a), and resulting antisense peptide pairings (b).
2. The second position/letter of the codon in the SGC table specifies physicochemical parameters of the amino acids—polarity, hydrophobic moment, lipophilicity and statistical potential.
3. Bases C or G at the second codon position code for the amino acid cluster with neutral property of the physicochemical parameter (2), while bases U and A at the second codon position code for the amino acid clusters with negative or positive property of the physicochemical parameter (2), respectively.
4. Peptides derived by means of genetic coding algorithm for the complementary DNA and RNA amino acid pairing interact with increased probability.

The interplay of rules 1, 2, 3 and 4 affects the selection and evolution of the protein and peptide ligand–receptor interactions. The Molecular Recognition Theory (MRT) relates complementary coding patterns of amino acid physicochemical properties to structural diversity, functional specificity and the evolutionary impact of protein and peptide ligand–acceptor interactions. We present and discuss a revisited Molecular Recognition Theory, considering the recently proposed Relational Model of the Standard Genetic Code [2].

References

[1] Štambuk N, Konjevoda P, Turčić P, et al. Genetic coding algorithm for sense and antisense peptide interactions. Biosystems, 164:199-216, 2018. doi: 10.1016/j.biosystems.2017.10.009.

[2] Konjevoda P, Štambuk N. Relational model of the standard genetic code. Biosystems, 210, 104529, 2021. doi: 10.1016/j.biosystems.2021.104529.

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

Mindmetacode, or gene expression and back again

Omar Paredes,
ElsaMagaña-Cuevas, Juan D. Serna-Grilló,Rebeca Romo-Vázquez, J. Alejandro Morales
Departamento de Bioingeniería Traslacional, CUCEI,
Universidad de Guadalajara, México
omar.paredes@alumnos.udg.mx

Codepoeisis is the essential process that enables organisms to transit into more complex functional states featuring new codes of life. According to Igamberdiev [1], living systems show a complexification phenomenon that involves embodying new structures without adding either molecules, regulations, or functions. This natural scheme entails that biological elements at different levels interconnect in a regulatory network where biological information flows through complex paths to link different biological levels. The regulatory networks encompass mesostructures or modules that yield emergent functions not feasible by stacking multiple organic codes. An example of this organization is outlined for cell differentiation, where the regulatory metacode maps from assemblies known as n-fold regulon tuples to a phenotypic continuum space[2]. The modularization of the regulatory interaction networks has yielded a complex architecture into an evolutive organ called brain [3]. In the brain, low-level organic codes along with next-level codes called neuronal codes merge into a metacode to achieve a hierarchical modular organization that outcomes into the emergence of early cognitive processes such as perception, representation, and memory. The above cognitive processes together with the metacode follow a complexification trend shaped by modularization and hierarchization to unfold higher-order features such as mind and consciousness. Such trend may set the outcome of diverse informational structures replicable inside and outside nature.

[1] Igamberdiev, A. U. (2021). The drawbridge of nature: Evolutionary complexification as a generation and novel interpretation of coding systems. *Biosystems*, *207*, 104454
[2] Paredes, O., Morales, J. A., Mendizabal, A. P., & Romo-Vázquez, R. (2021). Metacode: One code to rule them all. *Biosystems*, *208*, 104486.
[3] Paredes, O., López, J. B., Covantes-Osuna, C., Ocegueda-Hernández, V., Romo-Vázquez, R., & Morales, J. A. (2021). A Transcriptome Community-and-Module Approach of the Human Mesoconnectome. *Entropy*, *23*(8), 1031.

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

The Genetic Code – Between Biology and Computer Science

Paško Konjevoda and Nikola Štambuk
Ruđer Bošković Institute,
Bijenička cesta 54, 10000 Zagreb, Croatia
pkonjev@irb.hr

A popular approach in analysis of complex biological systems is their comparison with computers. The comparison may be applied to protein synthesis: 20 canonical amino acids can be considered as an alphabet that builds linear strings, i.e., proteins. In computer science terminology, the finite or infinite collection of strings is called a language. The rules for the construction of legal strings are called syntax. However, computer systems do not store individual letters in a direct textual format, instead they use a binary system which enables efficient storage of data. In order to use stored data, binary data must be converted into a textual form using a special code system called UNICODE [1]. This code system is completely comparable with the standard genetic code (SGC), with the difference that the SGC associates data stored as DNA triplets in DNA with amino acids.

This simple comparison helps to explain two important properties of the SGC—universality and the continuity principle of evolution. The construction of strings that are strictly based on characters coded by UNICODE requires only one program, that is, textual editor which concatenates characters in a linear order. In the same way, 20 amino acids coded by SGC can be linearly joined in a peptide/protein by a ribosome, an universal machinery for the construction of proteins. In this way an unlimited number of languages (e.g. biochemical pathways) can be devised by the same machinery and from the same alphabet. UNICODE evolved from a simpler coding system called ASCII. However, ASCII is completely embedded in UNICODE table, and in this way discontinuity between coding systems was avoided. All older textual information based on ACSII code can be interpreted by UNICODE without problem. The same continuity principle is preserved in SGC. Many properties of the 64 triplets that build SGC are inherited from simpler, ancestral codes based on 16 doublets, which are completely preserved in the SGC [2].

References

[1] Haddock SHD, Dunn CW. Practical computing for biologists. Sinauer Associates, Inc., Publishers Sunderland, Massachusetts U.S.A., 2011, Chapter 2, pp. 17-30.

[2] Konjevoda P, Štambuk N. Relational model of the standard genetic code. Biosystems, 210, 104529, 2021. doi: 10.1016/j.biosystems.2021.104529.

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

CODE BIOLOGY, MATRIX GENETICS AND THE DOCTRINE ON ENERGY-INFORMATION EVOLUTION BASED ON BIO-ANTENNA ARRAYS

Sergey V. Petoukhov
Mechanical Engineering Research Institute of the Russian Academy of Sciences
spetoukhov@gmail.com

The report presents the results of the author's research on modeling the informational characteristics of the molecular genetic coding system. These results are obtained using matrix analysis, quantum information formalisms and antenna array theory. It is shown that the system of oligomeric (n-plets) DNA alphabets is represented naturally as a tensor family of square matrices, the columns and rows of which are numbered by means of dyadic groups of binary numbers. The connection of these structured alphabets of DNA with algebraic bit-reversible holography, known in digital informatics, as well as with Poincaré's disk model of Lobachevsky's hyperbolic geometry is demonstrated. In light of this, the possible genetic origins of the links between genetically inherited physiological phenomena with hyperbolic geometry, described by different authors, are discussed. At the same time, the discussion includes the genetic foundations of physiological phenomena of the holographic type, which have long been described in physiology (K. Pribram, etc.). The important role of holographic models in modern physics (D. Bohm and others) is mentioned, as well as the widespread use of physical holographic methods in informatics due to their high usefulness in the tasks of processing and storing information. The revealed universal rules of stochastic organization of nucleotide DNA sequences in the genomes of higher and lower organisms and their algebraic features are highlighted [1, 2]. At the same time, genomic DNA appears as a set of many parallel texts, each of which is written on the basis of own n-plets alphabet. The findings draw attention to the fundamental problem of "probability-vs-determinism" in biology. It is shown that the stochastic organization of genomic DNA is a highly limited stochastic, in which the probabilistic characteristics are accompanied by the presence of many deterministic rules concerning the invariance of the summary probabilities in certain groups of n-plets. These data on genetic stochastics are discussed in connection with genetically inherited phenomena of Gestalt psychology and the author's concepts of Gestalt biology and Gestalt genetics. On the basis of the obtained research results, the author formulates and develops his doctrine of energy-informational biological evolution, using analogies of matrix genetics, quantum informatics and tensor-matrix theory of antenna arrays.

[1] Petoukhov S.V. Algebraic harmony and probabilities in genomes. Long-range coherence in quantum code biology. *Biosystems*, v. 209, 104503 (2021). https://doi.org/10.1016/j.biosystems.2021.104503.
[2] Petoukhov S.V. Algebraic Rules for the Percentage Composition of Oligomers in Genomes. - *Preprints 2021*, 2021010360, 3rd version, 84 pages (2021). DOI:10.20944/preprints202101.0360.v3

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

Circular codes and the SARS Corona Virus 2

Simone Giannerini and Diego Luis Gonzalez
University of Bologna, Department of Statistical Sciences
simone.giannerini@unibo.it
gonzalez@bo.imm.cnr.it

Alberto Danielli and Maria Van Eijk
University of Bologna, Department of Pharmacy and Biotechnology
alberto.danielli@unibo.it
maria.vaneijk@studio.unibo.it

Greta Goracci
Free University of Bolzano, Faculty of Economics
greta.goracci@unibz.it

Previous works showed that protein coding sequences possess statistical properties that can be linked to the theory of circular codes. Also, they found a link between translation accuracy and efficiency and circular code properties. In this paper we exploit such results to analyse the complete genome of the SARS-CoV2 that is characterized by a programmed frameshift and overlapping ORFs. Among other things, we show that the theory of circular codes is able to predict both the frameshift and the coding portion of the 5' leader sequence fused to genomic and subgenomic RNAs.

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

On the Minimal Elements of the Genetic Code – Again on the Structural Analogies between Phonemes and Nucleotides

Suren Zolyan
National Academy of Sciences, Republic of Armenia,Yerevan, Armenia
Immanuel Kant Baltic Federal University, Russia
surenzolyan@gmail.com

1. The comparison of the genetic code to language has been in place since the beginning of its discovery. The key terms of molecular genetics are based on metaphors describing biochemical reactions as operations on the text (reading, transcription, translation, proofreading, editing). Accordingly, the minimal element is the letter that denotes the nucleotides. Scientists even argue about which letter appeared first: G or C.
2. The equation of the basic units of the genetic code with the Latin alphabet letters is justified only as a conventional way of notation. Therefore, attempts to describe the genetic code as a semiotic system cannot be successful if the minimal unit of the genetic code is considered as an entity comparable to what a letter is in the system of language description (Cf.: Jakobson 1970, 438 ; Fatyneket al 2019). Letters, to be precise, are not elements of language but of culture, they were invented relatively late as a designation of sounds in writing. The minimal unit of language is sound/phoneme, and nucleotides can only be mapped to it. (cf. Zolyan 2021; Barbieri 2021;).
3. Meanwhile, the only one who likened nucleotides to phonemes was Fr. Jacob (1977), probably influenced by his interlocutor, Roman Jakobson, who made a very explicit clarification: "*Since our letters are mere substitutes for the phonemic pattern of language, and the Morse alphabet is but a secondary substitute for letters, the subunits of the genetic code are to be compared directly with phonemes*" (ibid). The successes of linguistics of the twentieth century were due to the distinction between language, system, and speech, that is, the manifestation of the system. First of all, it was a distinction between the abstract unit of the system, this is the phoneme, and the real sound. A phoneme is defined exclusively negatively as a minimal set of features that distinguish it from other phonemes.
4. Extrapolation of this methodology to the genetic code leads to a distinction between a nucleotide as a chemical element and a nucleotide as an intrasystem (abstract) unit of the genetic code. The latter is characterized by only two differential features - the number of hydrogen bonds and belonging to the group of purines or pyrimidines. Only these two features are essential for the processes of coding - differentiation of amino acids.
5. Analogies between phonemes and nucleotides as the minimum elements of the corresponding semiotic systems can be developed. Paradigmatic and syntagmatic regularities well described in phonological systems also take place in gene expression processes. Thus, like phonemes, neutralization occurs in weak positions - in the so-called degenerate codons, the third position is irrelevant, and in 32 cases can be filled with any nucleotide (neutralization of both features), and in 30 cases– with the nucleotide of the same group (neutralization of the number of hydrogen bonds). The same regularity is observed in syntagmatic relations: it is manifested as the unreadability of the third position. In linguistic terms, this is also explained as the neutralization of differential features. Such extrapolation allows us to offer a new interpretation to the phenomenon of wobbling (Crick 1966) and Lagerkvist's "two out of three" rule (Lagerkvist 1978): the codon - anticodon pair acts as a coherent translation unit, which is why the complementary pair of the 3rd position of a codon and 1st position of an anticodon turns out to be irrelevant (redundant).

References

Barbieri Marcello, Overview of the third special issue in code biology, Biosystems, Volume 210, 2021, 104553, https://doi.org/10.1016/j.biosystems.2021.104553.

Crick F. Codon-anticodon pairing: the wobble hypothesis // J. Mol. Biol. – 1966. – N 19. – P. 548–555

Faltýnek, D., Matlach, V. &Lacková, Ľ. Bases are Not Letters: On the Analogy between the Genetic Code and Natural Language by Sequence Analysis. Biosemiotics 12, 289–304 (2019). https://doi.org/10.1007/s12304-019-09353-z

Jacob F. The linguistic model in biology // Roman Jakobson. Echoes of his scholarship /. Eds. D. Armstrong, C.H. van Schooneveld. – Lisse: Peter de Ridder, 1977. – Pp.185-192

Jakobson, R. (1971). Linguistics in relation to other sciences. In Roman Jakobson, Selected Writings: Vol. 2:Word and Language, 655–696, The Hague — Paris: Mouton.

Lagerkvist U. «Two out of three»: An alternative method for codon reading // Proc. Natl. Acad. Sci. USA. – 1978. – N 75. – P. 1759–1762

Zolyan Suren, On the context-sensitive grammar of the genetic code, Biosystems, Volume 208, 2021,104497, https://doi.org/10.1016/j.biosystems.2021.104497

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

Synechism in the light of Code Biology. A bio-logical view of personhood

Valerio Marconi
Adjunct Professor of Philosophy of Language
Department of Humanities, University of Urbino Carlo Bo
valerio.marconi92@gmail.com

Despite their departure, Code Biology and Biosemiotics share a trivalent analysis of the sign, yet it can be shown that – contrarily to what is entailed by Peircean Biosemiotics – organic codes are not a special case of mental codes since their triadic structure can be reduced to dyadic relations (see V. Marconi, *Tra filosofia, semiotica e strutturalismo*, Treviso 2020: 241-263). Consequently, organic semiosis and mental semiosis are both trivalent processes, but only mental semiosis counts as an irreducibly triadic action. Thus, Peirce's synechism cannot be true: "all phenomena are of one character, though some are more mental and spontaneous, others more material and regular" (EP2: 2). Not all living matter is mental. Relying also on Giorgio Prodi's work, it is possible to sketch a restricted view of Peirce's synechism and to delimitate its scope to the cognitive systems of second and third type (see M. Barbieri, *Code Biology*, Cham 2015: 126-127), namely the minds of interpretative animals and humans. A first outcome of this conceptual operation is to introduce the hypothesis of a *personal code* such that personhood, i.e. a remarkable and unique feature of human mind, could be regarded as a peculiar case of cultural code. Prodi's Biosemiotics is much closer than Peircean approaches to biological methods while critically assessing and reshaping many Peircean ideas (see F. Cimatti, *A Biosemiotic Ontology*, Cham 2018). He wrote: "Were a man deprived of his [relational] term 'other man, network of others', this man would not simply be missing something: such a man does not exist […]. To sum up by means of a motto: love your neighbor since he is yourself" (G. Prodi, *Gli artifici della ragione*, Milano 1987: 167). From this, we can infer that personal identity does not follow different paths from cultural identity: personal connections between cultural objects and cultural meanings cannot be severed from social connections among different domains in cultural codes, so that each person is a community in constant exchange with other personal communities and wider interpersonal communities. Indeed, a consequence of Peirce's synechism seems to be true: "In the first place, your neighbors are, in a measure, yourself […]. In the second place, all men who resemble you and are in analogous circumstances are, in a measure, yourself" (EP2:2).

Empirical laws applied in biology

Vladimír Matlach
Palacky University, Olomouc, Czech Republic
vladimir.matlach@upol.cz

Concerning the genetic code, biology has a specific point in common with linguistics (and natural language processing): both fields can work with texts written in .txt files, using an alphabet of possibly 20-30 letters. Finding common topics and transferring knowledge from one field to the other, using the same tools (Markov Chains, Topic Modelling, LDA, etc.), measuring the same quantitative properties (such as Entropy), and using similar terminology is unsurprising. There is a topic of empirical laws seen and tested on natural languages, which are tested and observed at the molecular level of DNA. Such empirical laws have a few shared features: they are, by hypothesis, typically connected to the preservation and economization; they are also formally described by mathematical formulas but lack a definitive explanation. Both the genetic code and natural languages share some empirical laws. This contribution shows how problematic it is to assess such linguistic empirical laws on the genetic code, taking as an example of the Menzerath-Altmann Law (MAL). This law has been studied on several levels in the genetic code previously, most recently (by the author) on secondary structures of proteins. This resulted in a simple statement: as in natural languages with words and sentences, we find that for proteins a quantitative relation holds such that the more secondary structures (words) are used in a protein (sentence), the smaller the secondary structures become (measured in the number of letters or amino acids). Such a statement makes a full thought circle, from surprise through questions about its triviality, its reproducibility by stochastic processes, to end up again at the initial surprise. Such thought process and its results are the topics of this contribution, as they demonstrate the dangerous nature of empirical laws and their many traps when finally yielding implications for protein promoter regions, functions, and interactions.

Eighth International Conference in Code Biology – Olomouc (Czech Republic) 7-11 June 2022

Genetic and generic signs: referents, signifieds, and functions

Luccas Vaz Dantas dos Santos
Federal University of Parana, Brazil, Department of Philosophy
and
Wanderley Dantas dos Santos
State University of Maringá, Brazil, Department of Biochemistry
wdsantos@uem.br

In linguistics, signs have a non-referential domain where the signifier-signified relation is established (*langue*) and a discursive level where referents are defined. This linguistic split between the signified and the referent implies that the former has a broader utility that may serve specific purposes in the discursive context. Unlike in linguistic systems, genetic 'signs' establish a (more or less) bi-univocal relationship between signifiers (genetic information) and referents (amino acids, proteins and enzymatic products), without the mediation of 'general' meanings (signifieds). That is to say, the genetic referent is an immediate signified with only one context, i.e., it is indexical. This can be interpreted either as the unilateral relationship between a gene and a protein, or more broadly between a genotype and its phenotype. Throughout evolution, the genetic code undergoes a drift due to random mutations, being selected at the population level by the comparative advantages of its effects on the environment (competition, cooperation, phenotypic plasticity). This process of separation between an information recording domain (genetic system) and a domain of environment interactions (transcription and translation products of genetic information) is precisely what liberated biological systems from the determinism inherent in physical and chemical structures, including all other dissipative structures. This separation coincides with the emergence of a coding system capable of translating a set of states (a sequence of nucleotides) into a corresponding one (a sequence of amino acids). This process gave rise, at once, to the related properties of code, information, sign, signified and referent. These elements cannot exist apart. For instance, the translation system (tRNA-aatRNAt-ribosome) establishes a genetic code so that a sign in an mRNA (codon) translates to a specific amino acid, that is, its referent. Unlike the sign in Saussurian semiology that *internalizes* a relation between a signifier and a signified (avoiding considering the referent), the genetic sign requires a pre-modern (or pragmatic) semiotics wherein the referent is conceived as *external* to the sign. A genetic sign (codon) is reduced neither to quantity (i.e. Shannon's information) nor to a quality. It is a quality but only insofar as it correlates unilaterally with another quality (amino acid sequence and its derivatives). Nucleotides are measurable 'figures' (what the linguist Hjelmslev terms the meaningless components, such as phonemes or letters, that integrate linguistic signs), however they only gain relevance insofar as they combine into codons that correspond to specific amino acid sequences (polypeptides).Structures such as vortices, convection cells and flames do not rely on or even use information. Their maintenance stems from positive and negative feedback loops that stabilize in an energy gradient. This principle also applies to living beings that depend on an energy gradient to subsist. However, the organic and catalytic nature of living beings allowed the development of coding and replicating systems capable of evolving by natural selection. Thus, the primordial function of the code emerges as an organizing principle of the biological dissipative structure. To self-produce, the organism needs to recognize and select nutrients in the environment, a characteristic that makes genetic information somewhat a representation of the milieu (Umwelt). However, much of the information is focused on the perpetuation of the code itself: both in metabolism (semiopoiesis) and reproduction, as can be deduced from the universal set of genes, for example. Only secondarily does the genome begins to accumulate environmental information that enhances its performance in the ecosystem (e.g. chemotaxis, quorum sensing and so on).