

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

Issue: *DNA Habitats and Their RNA Inhabitants*

Life is physics and chemistry and communication

Guenther Witzany

Telos-Philosophische Praxis, Salzburg, Austria

Address for correspondence: Guenther Witzany, Telos-Philosophische Praxis, Vogelsangstrasse 18c, 5111-Buermoos, Austria. witzany@sbg.at

Manfred Eigen extended Erwin Schroedinger's concept of "life is physics and chemistry" through the introduction of information theory and cybernetic systems theory into "life is physics and chemistry and information." Based on this assumption, Eigen developed the concepts of quasispecies and hypercycles, which have been dominant in molecular biology and virology ever since. He insisted that the genetic code is not just used metaphorically: it represents a real natural language. However, the basics of scientific knowledge changed dramatically within the second half of the 20th century. Unfortunately, Eigen ignored the results of the philosophy of science discourse on essential features of natural languages and codes: a natural language or code emerges from populations of living agents that communicate. This contribution will look at some of the highlights of this historical development and the results relevant for biological theories about life.

Keywords: metaphysics; exact science; language; communicating agents

Introduction: from metaphysics to exact science

Today, everyone speaks about the genetic code—genes encoded in DNA that serve as the information-bearing molecules for all biological entities. Our main assumptions about the genetic code are inherently connected with leading scientists such as Erwin Chargaff, James Watson, and Francis Crick, who deciphered the rules and structure of its molecular syntax. In the 1970s, Manfred Eigen insisted that the genetic code really represents a natural language and is not just metaphorical.

Interestingly, at the same time, the discussion about the implementation of exact science, lasting several decades, led to clarification about the essential features of natural languages. A natural language or code emerges and is used by populations of living agents that communicate to organize and coordinate. If the genetic code really represents a natural language, we have to identify communication of competent code-using agents in addition to its physicochemical features.

We can explore the roots of the history of the discussion that finally led to biology becoming a subdiscipline of physics and chemistry within a

reductionistic paradigm. At the beginning of this history, the main themes were rather different.

After the principles of nature were explained by the successful means of the four elementary forces of water, air, fire, and earth, two fundamentally contradictory new paradigms (i.e., all is one versus all is many), dominated the ongoing history of philosophy in the last 2000 years. Nearly all philosophies and subsequent scientific theories were founded on one of these core paradigms, which largely represent the emotional presumptions and preferences derived from hidden myths and beliefs that we term metaphysics.

One metaphysical concept was holism. The ultimate conviction was that all beings are part of the one and only wholeness. This turned the classical antinomia "one" or "many" to "all is one." This means that there is one being, everything we wonder about, investigate, and use to modulate or take part in production processes is a chimera of the core material, the one energy flow, the substratum of everything. Proponents were Thales of Miletus, Heraclitus, Spinoza, Giordano Bruno, Leibniz, and Hegel, and more recently Huxley, Darwin, Spencer, Einstein, and Hawking. One law determines the whole universe. The one and only law is the

law of development with its everlasting potentials of differentiation and integration. Albert Einstein noted there was one main energy as in the first 10^{-43} seconds—the so-called big bang—and all atoms and subatomic particles were formed out of this, but the corpuscular particles we measure are solely condensations of this universal field of energy. The search by Stephen Hawking for the last and only formula is an example of a more recent adaptation of this kind of metaphysics.

The fundamental contradictory perspective is the metaphysical concept of the pluralistic mechanism “all is many.” There are indefinite little unchangeable particles. The closer we look at subatomic parts, the more we find little entities. All is constructed out of these parts in accordance with mechanistic natural laws. The dynamics of all parts could be fully predicted if one could oversee all of the present energy states. What we see are pure chimeras, because, in truth, all these phenomena are constructed out of these little modules. Proponents of this counter-paradigm include Parmenides and Democritus, and later Gassendi, Boyle, Proust, Dalton, Descartes, Newton, LaPlace, Maxwell, Boltzmann, and Perrin.

Because, in the last 2000 years, there have been endless discourses whether all is one or all is many, which cannot be decided rationally, a special kind of philosophy of science decided to avoid metaphysics in principle and to focus on exact science. From the 19th to the 20th century, therefore, philosophers and physicists such as Wittgenstein, Carnap, Neurath, Gödel, Russell, and Tarski suggested that exact sciences do not depend on myth and belief states but are strictly based on empirical facts coherent with observations and measurements in experimental setups. Logical empiricism should exclude metaphysical language, inexact terms, and apodictically proclaimed truth and only express empirical sensory data that can be expressed logically (i.e., by mathematical equations). This was a radical (linguistic) turn and a fundamental criticism of metaphysics.¹

From exact science to the fundamentals of natural languages/codes

The main assumptions of this linguistic turn are as follows. Metaphysical questions do not have any real subject and must be replaced by empirical scientific knowledge (i.e., materialism and naturalism). In science, there is an inner relationship

between thinking and being. The only value in science is the rationality of the methods of scientific knowledge, represented by the formalizable expression of empirical sentences.^{2–5} In strict objectivism, the pure observer confirms observations through measurement techniques and subsumes reality in the formalizable depiction of these measurements. The core thinker in this context was the young Ludwig Wittgenstein: in his view, every sentence with which we describe observations as well as sentences that are used to construct theories must fulfill the criterion of formalizability. Because the world functions exclusively according to the laws of physics, it can be depicted only by mathematical sentences able to depict physical reality in a one-to-one manner. Natural laws expressed within the language of mathematics represent the inherent logic of material reality. The most important element of language is its syntax, because only a logical syntactic structure can depict the logical structure of material reality. Behind this logic of syntax that represents the logic of material reality, there is a universal grammar that is in coherence with the laws of physics.⁶ Sentences out of this realm are scientifically senseless, with the consequence “Whereof one cannot speak, thereof one must be silent.”⁷ This radically reduced exact science to scientific sentences, which are formalizable (i.e., can be expressed as mathematical equations).

In addition to these attempts, cybernetic systems theory and information theory were developed. Both disciplines investigate the empirical significance of scientific sentences using a quantifiable set of signs. They investigated the transfer of formalized references between a sender and a receiver: understanding information is possible because of the logical structure of the one and only universal syntax, through a process that reverses the construction of meaning. The syntactic structure of a scientific language can be expressed in binary code (1/0). Therefore, it can be sufficiently described and investigated by computers with algorithm-based programs.^{8–12}

The astonishing shift in argumentation led to the investigation of the function of natural languages. If we want to define exact sciences, we have to define what a sentence has to fulfill if it wants to contribute: we have to identify the essential criteria that must be fulfilled if we are to term a sentence scientifically correct. Then, we can clearly delimit metaphysical statements from scientific sentences.

Where and how do natural languages evolve? As we are humans, we must find descriptions of our own natural language that are not only true and correct in a scientific sense but, additionally, agree with the results intuitively and evidently. How do we achieve this?

Two earthquakes

The rise of mathematics as the supreme discipline of science at the beginning of the 20th century culminated in David Hilbert's universal program for constructing an axiomatic system with error-free logical sentences. This was outlined in a broader realm of foundation and justification by Bertrand Russell and Alfred North Whitehead.^{13,14} Kurt Gödel's incompleteness theorem fundamentally destroyed these attempts.¹⁵ Gödel showed that, in any thinkable axiomatic system, there is always, in principle, the possibility of constructing really new sentences that cannot be deduced from existing ones, that cannot be predicted, and for which no algorithm is available. The construction of an axiomatic system with error-free logical sentences is impossible in principle.

In addition to this "earthquake" in the foundations of exact science, Ludwig Wittgenstein rebutted, in his later philosophy, the theoretical construction of a universal world-depicting and formalizable language.¹⁶ Before we are able to learn and construct artificial languages such as those used in scientific disciplines, and even before we can know what we should investigate or what goals of research should look like, we learn everyday language. Everyday language is the ultimate precondition of our cognitive competence, and commonly shared understanding of utterances is the ultimate prerequisite of common coordination of behavior. It is not possible to go behind this everyday language, as it is not possible to formulate or even think about sentences that look at language, being, or material reality from a position outside language.

Results: essential features of natural languages/codes

At this stage, a new paradigm arose out of the investigations on natural languages. Natural languages are essential tools of populations of living agents for commonly coordinating behavior in that they share combinatorial rules (syntax), content-

specific rules (semantics), and context-dependent rules (pragmatics).

Because natural languages depend on populations in which individuals have to integrate into group identity, which means having to coordinate behavior, language use is principally a kind of social interaction. It depends on two competences. First, a communicative competence for correctly installing interactions between at least two agents. Second, a linguistic competence for generating correct sign sequences.¹⁷

Natural languages do not speak themselves, but depend on living agents that are able to combine signs correctly to form more complex sign sequences, to install and fulfill social interactions, and therefore to represent the content of information relevant for interactional motifs. Natural languages serve to install social interactions and rely on pragmatic, syntactic, and semantic rules. These rules are rather conservative, but, in contrast to natural laws, may change according to circumstance (adaptational purposes).

The crucial difference to the former paradigm of exact science thinking was that pragmatics became primary, not syntax, because the pragmatic real-life context is of crucial relevance for the meaning of signs. After the linguistic turn, the pragmatic turn was introduced into the philosophy of science discourse with late Wittgenstein, Austin, and Searle,^{18,19} because it became self-evident that not the syntax but the context (pragmatics) determines meaning: identical sign sequences can represent different and even contradictory meanings in different situations (e.g., if we look at the phrase "The shooting of the hunters"; Fig. S1).

Has a pragmatic turn occurred in molecular biology, genetics, and RNA biology?

No! Since Erwin Schrödinger's "life is physics and chemistry,"²¹ no other approach has integrated current knowledge about the functions of natural languages/codes. The rise of information through cybernetic system theory reduced cells, organs, and organisms to systems that emerge like self-reproducing automata similarly to information theory.^{8,12,22} Both served as theoretical springboards from which to investigate nucleic acid sequences as a quantifiable set of signs similar to

information-processing systems and subsystems. Someone has to detect the logical structure of the universal syntax hidden in the sequence structures to know what its meaning is (a curious presumption that underpinned the Human Genome Project that intended to identify all existing genes within the human genome out of the sequence syntax).

Manfred Eigen integrated these newly emerging mathematical theories of language in his quasispecies theory and theory of hypercycles in the early 1970s. Evolutionary novelty is the result of replication errors of the fittest types. A mass of such replication errors in RNA groups such as viruses produces varieties that compete with the fittest master type.^{23–25} These variations are the core material of biological selection.

Human brains are the result of evolutionary selection processes that favor the physical assembly of neuronal cells based on physics and chemistry. The best depiction that brains can obtain from material reality is its sensory program and its description by formalizable sentences (i.e., mathematical equations). Formalizable sentences are the only appropriate language for the brain to depict itself in its material reality.²⁶

For Eigen, both the logic of the describing and that of the theoretical constructing language corresponded to the logic of the system. The relationship between these elements can be represented in an abstract, formal, and unambitious manner. This means that the best investigation of quantitative measurements can be done by algorithm-based computing machines. Eigen could not conceive of *de novo* generation of nucleic acid sequences and sentences of human language without algorithm-based processing. He did not realize the consequences of the discussions about the features of natural languages and codes and retained this old mathematical theory of language in which the meaning of sign sequences is determined by a (theoretically suggested) universal grammar inherent in material reality.²⁷ In his view, the interactions of RNAs in hypercycles, as well as the interactions of humans, underlie the logic of material reality. Therefore, social behavior also remains a feature of material reality, and can best be investigated by mathematics and computers.

How do we adapt the pragmatic turn to RNA biology and virus-first scenarios?

Eigen's quasispecies and hypercycle concepts have dominated evolutionary biology for 40 years, with only marginal changes occurring. But if mathematical, systems theoretical, and information theoretical concepts can investigate only quantifiable aspects of the genetic code, we can ask how to identify qualitative aspects such as competences of agents that edit the genetic code.²⁸

If we look at current research in RNA biology, we can identify such qualitative aspects. Increasing knowledge about the abundance of non-coding RNAs and their function in regulation of gene expression, transcription, RNA processing, translation, immune functions, DNA synthesis, genome maintenance, rearrangement, and repair show multiple motifs of behavior that are qualitatively ruled.^{29–33}

If we consider at the nucleic acid sequences to be natural code, we need to identify real-life communication processes as social interactions. This means that we have to identify competent agents that act as consortia and edit genetic code according to their selfish purposes or, more successfully, co-opt and exapt it out of former infectious RNA agents, for both their successful group behavior and the purposes of their host organisms.³⁴ Additionally, these agents must share some common semiotic (pragmatic, syntactic, semantic) rules in genetic content composition and generation of *de novo* sequences according to context-dependent needs such as elementary stress situations, changing environmental conditions, or even radiation or chemical influences.³⁵

The field of research that offers such competent agents involves the abundance of noncoding RNAs. Most noncoding RNAs operate in complex with proteins as ribonucleoprotein particles such as ribosomal subunits, spliceosomes, editosomes, small nuclear and nucleolar RNPs, microRNAs, long noncoding RNAs, and the ancient and fascinating world of reverse transcriptases.^{36–39} Many noncoding RNAs are able to base pair to other nucleic acids, especially the non-base-pairing loops of the RNA stem loops, which are rather active in interacting processes. Interestingly, both “kissing-loop” and pseudoknot RNA biology demonstrate convincingly

that the meaning/function of such consortia depend on context, not on sequence syntax.^{39,40}

A variety of concepts deals with an active form of life before cellular life emerged.^{41,42} The role of viruses and consortia-generating ribozymes is also a serious and far-reaching consideration, because it represents a variety of indicators of the virus-first perspective in that viruses are the first RNA stem loop group-derived agents that predated cellular DNA-based life. How does this fit into our picture of genetic codes and life?

Are replication errors the driving force for variation?

In a purely quantitative paradigm, as represented by the exact-science thinking, emergence of information must be justified as a random process in which selective forces outcompete less significant information entities. If information is really only physics and chemistry, then replication can only form identical copies. If changes occur that vary the master copy, the correct designation is *error*. Therefore, mutation as error replication is the appropriate term, because it is a statistical event, randomly derived. In an abundance of replication errors, in very few cases it may occur that one or a few are beneficial to prior master types. They have selective advantages, and the accumulation of such rare beneficial mutations is the driving force for variation, the core force that is subject to biological selection.^{23–26}

Some decades ago, it became obvious that DNA is not an unchangeable information carrier but is subject to dynamic interactional change that has essential relevance for evolutionary novelty.^{43–46} It has been found that, in some events, a duplication of genes happened. After a duplication, when a copy is usually released from selective pressure, there might be rapid changes that either lead to gene inactivation or, in rarer cases, exaptation into a new or variant function.⁴⁷ However, this concept does not explain generative competences for evolutionary novelty in terms of genome editing.

Without infection no evolution: examples

Let us look briefly at three examples in which evolutionary novelty is not the result of error replication in an assembly of small single steps or, as in the duplication scenario, as a whole quantitative increasing rate of mutations, but clearly represents contradictory motifs.

First, there is the success story of *sympyogenesis* theory, which replaces the evolution of eukaryotes dependent on error replications. The serial endosymbiotic theory was revolutionary because it changed the perspective on eukaryotic evolution from accumulations of random mutations to social coordination of formerly free-living prokaryotes.⁴⁸ Eukaryotic cells are the result of the merging of several different ancestral genomes. Endosymbiotic concepts demonstrated convincingly that most of the DNA found in the cytoplasm of protoctists, animals, fungi, and plants come from genes of bacteria that became organelles, and not from mutations. Eukaryotic genes that participate in information processing (translation, transcription, etc.) show a close relationship to eubacteria. On the other side, genetic factors that control metabolic processes more closely resemble the metabolic processes in archaea. Additionally, several capabilities of the eukaryotic nucleus show features common to large DNA viruses and cannot be found in any cellular organism, which may indicate that the sympyogenesis of eukaryotes additionally depended on a persistent virus as its nucleus and maybe even its nucleolus.⁴⁹

If we were to investigate the evolutionary novelty of placental mammals, we would have to explain a rather complex genotype, the syncytin complex, by randomly derived replication errors. However, generation, integration, and regulation of the placenta are not a stepwise process, and missing links have not yet been found. A coherent explanation model would be as successful as any other explanation of complex organs. Today, we know that it was a strong infection event with persistent viral settlers that transferred complete genetic blocks into single or very few infection events to host organisms that changed the phenotypic organism in a radical way without an infinite number of small steps. The syncytin genotype enables the mother organism's immune system to not detect the embryonal genotype, which has some nonself genetic identities. The mother's immune system, therefore, cannot reject the embryo until a certain stage of development. The syncytin gene complex was transferred via a persistent retroviral infection event and not by mutations of previously available genetic sequences.⁵⁰

A similar evolutionary event is the evolution of *Homo sapiens*. Generations of scientists have wondered which assembly of replication errors led to the evolution of large social brain organs with their

key competence of human language. Despite the fact that vocal sound systems followed gestural systems, the reason for the evolution of large social brains remained unknown. Now, we have an alternative scenario that sounds rather speculative: massive infection events by endogenous retroviruses replaced or damaged the primary self/nonself differentiation capability in great apes, the vulnonaseral complex. This means that persistent retroviral integration damaged the whole identification system of values for great ape populations. This led to alternative self/nonself differentiation competences, especially in the mother–child relationship. Maternal care and social bonding had to be based on other capabilities such as auditory faculties (vocal sign systems) in combination with visual differentiation. A dramatic increase of visual and auditory competences would necessarily have been the result.⁵¹

Context, not syntax, determines meaning: epigenetic imprinting codes environmental needs

One of the main deficits of information theory and systems theory, introduced by Manfred Eigen into molecular biology, is their syntax centrism. Syntax-based quantitative approaches that focus on nucleic acid sequences as information-bearing molecules will have problems explaining why identical nucleic sequences may have different and, in extreme cases, even contradictory meanings. This energy-saving technique is made possible through epigenetic imprinting. The marking of nucleic acid sequences at the genome determines meaning—the function for which the sequences code. This may lead to a rather dynamic silencing or amplification of certain regions, depending on the real-life organismal circumstances. In particular, the infection-derived mobile genetic elements are possible candidates for epigenetic control and serve as the interface between experienced real-life events and genome change.^{52–54} For example, we know that plants can reactivate the genetic features of their grand- or great-grandparents if the genetic content they inherited from their parents does not have the appropriate competence to resist environmental circumstances.⁵⁵

Epigenetic marking signifies a crucial ability of living agents: memory and learning. It is a technique for imprinting the experiences of organisms of, for example, environmental experiences, stress

situations, or other circumstances.⁵⁶ First, it was thought that only humans could memorize and learn, but subsequently higher mammals such as great apes, elephants, and cetaceans were also believed to have this facility. At the current stage of research on RNA-mediated epigenetic processes, we know that, even in prokaryotes, epigenetic imprinting exists, and that bacterial populations can memorize and learn to adapt faster and tailor their behavior to special situations/environmental circumstances.

RNA sociology may explain qualitative motifs of RNA group behavior

The question remains how to define nucleic acid sequences as a real natural language/code, which means we have to identify competent agents that edit this code and, additionally, delete the error-replication narrative to coherently explain the generation of genetic novelty. Within the last decade, increasing knowledge about the abundance of non-coding RNAs that play important roles in all cellular functions, especially in genetic- and epigenetic-regulation processes, has replaced the dominant interest in the genetic storage medium DNA.^{57,58} Now, RNA is in the driver's seat. Investigations on the abundance of RNA functions shed new light on how cellular life is organized and coordinated.^{59,60} Additionally, RNA viruses that persistently infect cellular life via self-counterbalancing addiction modules, such as toxin/antitoxin, restriction/modification, and endonuclease/ligation modules, show modular co-opted exaptations of viral parts for cellular needs such as LINES, SINES, Alus, all the long terminal repeats (LTRs) and non-LTRs, and the abundance of small and large noncoding RNAs found in cellular regulation.^{61,62}

In contrast to the quantifiable analysis of molecular biology in RNA biology, there are several features that represent qualitative motifs of behavior that can better be subjected to a RNA sociology (Fig. 1).

- Acting agents of natural codes can generate *de novo* sequences and behavioral motifs that represent novel sequences that cannot be described sufficiently by mathematical theory of language-derived approaches because they can generate sign sequences for which no algorithm is available in principle.

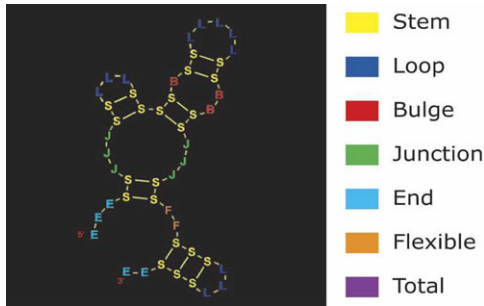


Figure 1. The RNA stem loops have several distinct parts/subunits: stems consisting of base-paired nucleic acids and loops/bulges/junctions consisting of unpaired regions limited by stems. Importantly, any RNA is part of such stem loops. Interestingly, single-RNA stem loops exclusively interact according to physical chemical laws only, whereas RNA stem loop groups follow biological selection. Additionally, it was found that cooperative RNA stem loops outcompete selfish ones. From Ref. 64.

- In contrast, a qualitative RNA sociology emphasizes the understanding of social interactions through analysis of active RNA sequences according to contextual needs. This means that pragmatic interactional motifs are crucial for meaning functions before sequence syntax.
- The exaptations of former (infection-derived) inventions to more appropriate needs fit very well into these investigations, offering a highly dynamic modular perspective on these interacting agents.
- Qualitative RNA sociology investigates the roles of RNA stem loops in the RNA stem loop–group behavior as social interactions within a perspective that focuses on group identity, self/nonself differentiation competence, quasispecies consortia in dynamic states, and long-lasting highly conserved stages.
- Additionally, such an RNA sociology investigates the behavioral rules that, unlike underlying natural laws, may change according to contextual needs such as environmental changes.
- Because, in this perspective, crucial parameters are the social character of interactions (not individual fittest types), and owing to the primacy of context rather than syntax structure, the role of group membership for group identities is as important as the remaining single-RNA stem loops that are outcompeted and degraded: they may remain important informational subunits that, in later evolu-

tionary and developmental stages, can be integrated again in a modular manner in other contexts with co-opted functions for other needs. This means that outcompeted single-RNA stem loops remain as important resources.

The explanatory force of a qualitative RNA sociology for RNA biology would be a more coherent explanation of both RNA social interactions and natural genome editing,⁶³ which includes RNA social interactions such as dynamics from single-RNA stem loops to group (identity) building, self-/nonself identification competence, context-dependent varying interactions, cooperation between RNA groups and host, *de novo* generation of nucleic acid sequences, coherent integration into preexisting ones, innovation by variations in the RNA stem loops, and innovative genetic identities formed by coevolution. In this respect, a qualitative RNA sociology may solve the basic problem of adequate explanation of RNA group behavior by adapting its methodological foundations to a pragmatic turn.⁶⁵

Conclusions

It is possible to extend the quantitative paradigm with its physicochemical core assumptions to find more data on RNA folding, the discovery of new RNAs with algorithms that couple thermodynamics with chemical mapping, and sequence comparison. This would update past and recent research results and enrich big-data files. It is also possible to measure the exchange of oxygen molecules in the airstreams produced by people speaking to each other, and vocal comparisons will detect syntactic structures and grammar following alphabetical orders that contain information. But however perfectly this quantitative approach may be realized, it will not understand the behavioral motifs of living agents that communicate by using a natural language. The follow-up of signals within signal sequences does not represent the context in which such sequences are used.

Additionally, the most intriguing competence of living agents is their capability for *de novo* generation of sequences, which does not occur as a replication error event. *De novo* generation of sequences in natural languages and codes is essential if communicating agents are to coordinate or to

invent new pathways of problem-solving strategies in real-life contexts. For *de novo* generation of coherent content in natural languages/codes, no algorithms are available; in other words, it is not computable. The examination of the history of this discourse showed that communication of living agents is amenable to formalizable procedures only to a limited extent, such as comparison of previous quantitative data, and cannot be reduced to the laws of physics and chemistry solely. Looking at nucleic acid sequences from the perspective that DNA serves as a preferred habitat for communicating RNA inhabitants could provide an inspiring exercise for biology, ridding it of its role as subdiscipline of physics and chemistry.

Conflicts of interest

The author declares no conflicts of interest.

Supporting Information

Additional supporting information may be found in the online version of this article.

Figure S1. In natural languages/codes, the meaning of syntactical identical sequences depends on the real-world context. Natural languages/codes can transport a superficial (visible) grammar and an invisible deep grammar (which transports the intentions of the sign user), which is marked by, for example, emphasis and facial expression. In contrast to the conviction of Sydney Brenner, algorithm-based machines (computers) that must extract the meaning of given syntax structures cannot differentiate between superficial grammar and deep grammar in principle. From Ref. 20.

References

1. Witzany, G. 2010. *Biocommunication and Natural Genome Editing*. Dordrecht, the Netherlands: Springer.
2. Carnap, R. 1931. Überwindung der Metaphysik durch logische Analyse der Sprache. *Erkenntnis* 2: 219–241.
3. Carnap, R. 1931. Die physikalische Sprache als Universal-sprache der Wissenschaft. *Erkenntnis* 2: 432–465.
4. Carnap, R. 1934. *Logische Syntax der Sprache*. Vienna: J. Springer.
5. Carnap, R. 1939. *Foundations of Logic and Mathematics*. Chicago, IL: University of Chicago Press.
6. Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
7. Wittgenstein, L. 1922. *Tractatus Logico-Philosophicus*. London: Kegan Paul.
8. Bertalanffy, L. 1940. Der Organismus als physikalisches System betrachtet. *Die Naturwissenschaften*. 28: 521–531.
9. Wiener, N. 1948. *Cybernetics, or Control and Communication in the Animal and the Machine*. New York, NY: Wiley.
10. Shannon, C.E. & W. Weaver. 1949. *The Mathematical Theory of Communication*. Urbana: University of Illinois Press.
11. Turing, A. 1950. Computing machinery and intelligence. *Mind* 59: 433–460.
12. Neumann, J. 1966. *Theory of Self-Reproducing Automata*. Urbana/London: University of Illinois Press.
13. Hilbert, D. & P. Bernays. 1934/1939. *Grundlagen der Mathematik*. Berlin, New York: Springer.
14. Whitehead, A.N. & B. Russell. 1910/1912/1913. *Principia Mathematica*. Cambridge: Cambridge University Press.
15. Gödel, K. 1931. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme. *Monatshefte für Mathematik und Physik* 38: 173–198.
16. Wittgenstein, L. 1953. *Philosophical Investigations*. Oxford: Basil Blackwell.
17. Habermas, J. 1994. Actions, speech acts, linguistically mediated interactions and the lifeworld. *Philosophical Problems Today* 1: 45–74.
18. Austin, J. 1962. *How to Do Things with Words*. Oxford: Clarendon Press.
19. Searle, J. 1969. *Speech Acts: An Essay in the Philosophy of Language*. London: Cambridge University Press.
20. Witzany, G. & F. Baluška. 2012. Life's code script does not code itself. The machine metaphor for living organisms is outdated. *EMBO Rep.* 13: 1054–1056.
21. Schroedinger, E. 1944. *What is Life? The Physical Aspect of the Living Cell*. London: Cambridge University Press.
22. Brenner, S. 2012. Turing centenary: life's code script. *Nature* 482: 461.
23. Eigen, M. 1971. Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* 58: 465–523.
24. Eigen, M. 1993. The origin of genetic information: viruses as models. *Gene* 135: 37–47.
25. Eigen, M. 2013. *From Strange Simplicity to Complex Familiarity: A Treatise on Matter, Information, Life and Thought*. Oxford: Oxford University Press.
26. Eigen, M. & R. Winkler. 1975. *Das Spiel: Naturgesetze steuern den Zufall*. München: Piper.
27. Witzany, G. 1995. From the “logic of the molecular syntax” to molecular pragmatism. *Ev. Cog.* 1: 148–168.
28. Witzany, G. (Ed.). 2010. *Natural Genetic Engineering and Natural Genome Editing*. New York, NY: Wiley & Blackwell.
29. Brosius, J. 1999. RNAs from all categories generate retrosequences that may be exapted as novel genes or regulatory elements. *Gene* 238: 115–134.
30. Sciamanna, I., P. Vitullo, A. Curatolo, et al. 2009. Retrotransposons, reverse transcriptase and the genesis of new genetic information. *Gene* 448: 180–186.
31. Mattick, J.S. 2009. Deconstructing the dogma: a new view of the evolution and genetic programming of complex organisms. *Ann. N. Y. Acad. Sci.* 1178: 29–46.

32. Clark, M.B., A. Choudhary, A. Smith, *et al.* 2013. The dark matter rises: the expanding world of regulatory RNAs. *Essays Biochem.* **54**: 1–16.
33. Mercer, T.R. & J.S. Mattick. 2013. Understanding the regulatory and transcriptional complexity of the genome through structure. *Genome Res.* **23**: 1081–1088.
34. Witzany, G. 2009. Noncoding RNAs: persistent viral agents as modular tools for cellular needs. *Ann. N. Y. Acad. Sci.* **1178**: 244–267.
35. Witzany, G. 2012. “From molecular entities to competent agents: viral infection-derived consortia act as natural genetic engineers.” In *Viruses: Essential Agents of Life*. G. Witzany, Ed.: 407–419. Dordrecht, the Netherlands: Springer.
36. Witzany, G. 2011. The agents of natural genome editing. *J. Mol. Cell Biol.* **3**: 181–189.
37. Witzany, G. 2008. The viral origins of telomeres, telomerases and their important role in eukaryogenesis and genome maintenance. *Biosemiotics* **2**: 191–206.
38. Teng, S.C., B. Kim & A. Gabriel. 1996. Retrotransposon reverse transcriptase-mediated repair of chromosomal breaks. *Nature (Lond.)* **383**: 641–644.
39. Flores, R., P. Serra, S. Minoia, F. Di Serio & B. Navarro. 2012. Viroids: from genotype to phenotype just relying on RNA sequence and structural motifs. *Front. Microbiol.* **3**: 217. doi: 10.3389/fmicb.2012.00217
40. Staple, D.W. & S.E. Butcher. 2005. Pseudoknots: RNA structures with diverse functions. *PLoS Biol.* **3**: e213.
41. Villarreal, L. 2005: *Viruses and the Evolution of Life*. Washington, DC: ASM Press.
42. Koonin, E.V. 2009. On the origin of cells and viruses: primordial virus world scenario. *Ann. N. Y. Acad. Sci.* **1178**: 47–64.
43. Shapiro, J.A. 1983. “Variation as a genetic engineering process.” In *Evolution from Molecules to Men*. D.S. Bendall, Ed.: 253–270. Cambridge: Cambridge University Press.
44. Shapiro, J.A. 1995. Adaptive mutation: who’s really in the garden? *Science* **268**: 373–374.
45. Shapiro, J.A. 2009. Revisiting the central dogma in the 21st Century. *Ann. N. Y. Acad. Sci.* **1178**: 6–28.
46. Shapiro, J.A. 2011. *Evolution: A View from the 21st Century*. Washington, DC: FT Press.
47. Zhang, J. 2003. Evolution by gene duplication: an update. *Trends Ecol. Evol.* **18**: 292–298.
48. Margulis, L. 1970. *Origin of Eukaryotic Cells*. Yale: University Press.
49. Bell, P. 2001. Viral eukaryogenesis: was the ancestor of the nucleus a complex DNA virus? *J. Mol. Evol.* **53**: 251–256.
50. Pérot, P., P.A. Bolze & F. Mallet. 2012. “From viruses to genes: syncytins.” In *Viruses: Essential Agents of Life*. G. Witzany, Ed.: 325–361. Dordrecht, the Netherlands: Springer.
51. Villarreal, L.P. 2009. *Origin of Group Identity: Viruses, Addiction and Cooperation*. New York, NY: Springer.
52. Shapiro, J.A. 2004. A 21st century view of evolution: genome system architecture, repetitive DNA, and natural genetic engineering. *Gene* **345**: 91–100.
53. Shapiro, J.A. 2013. How life changes itself: the Read-Write (RW) genome. *Phys. Life Rev.* **10**: 287–323.
54. Shapiro, J.A. 2014. Epigenetic control of mobile DNA as an interface between experience and genome change. *Frontiers Genet.* **5**: 1–16
55. Lolle, S.J., J.L. Victor, J.M. Young, *et al.* 2005. Genome wide non-Mendelian inheritance of extra genomic information in Arabidopsis. *Nature* **434**: 505–509.
56. Mattick, J.S. 2010. RNA as the substrate for epigenome–environment interactions: RNA guidance of epigenetic processes and the expansion of RNA editing in animals underpins development, phenotypic plasticity, learning, and cognition. *BioEssays* **32**: 548–552.
57. Atkins, J.F, R.F. Gesteland & T. Cech (Eds.). 2010. *RNA Worlds. From Life’s Origins to Diversity in Gene Regulation*. New York, NY: Cold Spring Harbor Laboratory Press.
58. Cech, T.R. & J.A. Steitz. 2014. The noncoding RNA revolution—trashing old rules to forge new ones. *Cell* **157**: 77–94.
59. Morris, K.V. & J.S. Mattick. 2014. The rise of regulatory RNA. *Nat. Rev. Genet.* **15**: 423–437.
60. Brosius, J. 2014. The persistent contributions of RNA to eukaryotic gen(ome)e architecture and cellular function. *Cold Spring Harb. Perspect. Biol.* 2014. 6. pii: a016089. doi: 10.1101/cshperspect.a016089.
61. Villarreal, L.P. 2012. “The addiction module as a social force.” In *Viruses: Essential Agents of Life*. G. Witzany, Ed.: 107–145. Dordrecht, the Netherlands: Springer.
62. Mruk, I. & I. Kobayashi. 2014. To be or not to be: regulation of restriction-modification systems and other toxin-antitoxin systems. *Nucleic Acids Res.* **42**: 70–86.
63. Villarreal, L.P. & G. Witzany. 2013. The DNA Habitat and its RNA Inhabitants: at the dawn of RNA sociology. *Genom. Ins.* **6**: 1–12.
64. Smit, S., M. Yarus & R. Knight. 2006. Natural selection is not required to explain universal compositional patterns in rRNA secondary structure categories. *RNA* **12**: 1–14.
65. Witzany, G. 2014. Pragmatic turn in biology: from biological molecules to genetic content operators. *World J. Biol. Chem.* **5**: 279–285.