Artificial and Natural Genetic Information Processing

Guenther Witzany

Telos-Philosophische Praxis
Vogelsangstr. 18c, 5111-Buermoos, Austria
witzany@sbg.at

Conventional methods of genetic engineering and more recent genome editing techniques focus on identifying genetic target sequences for manipulation. This is a result of historical concept of the gene which was also the main assumption of the ENCODE project designed to identify all functional elements in the human genome sequence. However, the theoretical core concept changed dramatically. The old concept of genetic sequences which can be assembled and manipulated like molecular bricks has problems in explaining the natural genome-editing competences of viruses and RNA consortia that are able to insert or delete, combine and recombine genetic sequences more precisely than random-like into cellular host organisms according to adaptational needs or even generate sequences de novo. Increasing knowledge about natural genome editing questions the traditional narrative of mutations (error replications) as essential for generating genetic diversity and genetic content arrangements in biological systems. This may have far-reaching consequences for our understanding of artificial genome editing.

Keywords: Genetic code; genetic engineering; natural code editing; RNA consortia; viruses.

1. Introduction

The dominating concepts of molecular biology and genetics in the last half a century, (i) the one gene-one protein hypothesis, (ii) the central dogma of molecular biology (DNA-RNA-protein-anything else), and (iii) the assumption that noncoding DNA is 'junk', are falsified meanwhile [Shapiro, 2009, 2011]. Since the rise of epigenetics the focus on the logic of molecular syntax of genetic sequences has lost its importance, because methylation and histone markings may add multiple meaning functions to identical molecular sequence syntax [Slotkin and Martienssen, 2007; Jirtle, 2009; Barlow, 2011]. Also it is becoming increasingly clear that noncoding RNAs serve as key regulatory elements in all steps and even sub-steps of replication, transcription, translation, recombination, repair and immunity [Mattick, 2009; Witzany, 2009; Cech and Steitz, 2014]. Most interestingly, research on the roles of persistent viruses in host genomes as main drivers of evolutionary processes, their various roles as mobile genetic elements, and their remaining roles as 'defectives' integrated as counterbalanced modules such as, e.g. restriction/modification, insertion/deletion and toxin/antitoxins, shows the abundance of agents competent in terms of arranging genetic content by integrating persistently into host genomes without destroying former coding regions [Villarreal, 2005, 2009; Koonin, 2009, Mruk and Kobayashi, 2014; Curcio and Derbyshire, 2003]. How does this current empirical knowledge fit the old core assumption of molecular biology and genetics and its theoretical concepts of the genetic code? This review will highlight some historical perspectives and compare them with the recent advances in the understanding of natural genome editing.

2. The Detection of the Genetic Code, Artificial Genetic Engineering and Genome Manipulation

Soon after the detection of the molecular syntax of the genetic code and its molecular biological features the idea arose of technically manipulating genetic content arrangements for various goals such as optimisation of the human gene pool, fighting various diseases, knocking out dangerous genes, optimising plant and animal breeding, and developing new techniques such as gene therapy. 'Now, biological research is in a ferment, creating and promising methods of interference with "natural processes" which could destroy or could transform nearly every aspect of human life which we value' [Wolstenholme, 1963]. Finally, the hope is that if the human genome can be deciphered completely then it will be easier to fight the major diseases affecting humans.

At the dawn of artificial genetic engineering mutations were caused, e.g. in plants, with few expected beneficial results at the beginnings that were object to further breedings. In a second step the real history of artificial genetic engineering began, with the manipulation of restriction enzymes, recombination of DNA in bacteria, better sequencing methods, and polymerase chain reaction. Interestingly, at this stage it was the investigation, understanding and use of virus-derived capabilities represented by plasmids and phages which were technically exploited. However, genetic engineering was thought not only to recombine genetic content arrangements within one species but also to apply a transspecies method to develop multiresistant plants, new drugs and even gene therapy. This was the consequence of the realisation that the genetic code is used by all living entities on this planet.

New insights into DNA splicing and the rise of epigenetics made it increasingly clear, however, that the molecular syntax of the genetic storage medium DNA did not really represent what is finally transcribed into RNA and translated into proteins, which means different epigenetic marking of identical genetic sequences could lead to different and in extreme cases opposing protein coding functions [Mattick and Gagen, 2001; Mattick, 2010; Tang

et al., 2015, Werner *et al.*, 2015]. Yet it is clear that the epigenetic markings on the genome are of similar importance to the sequence syntax. Epigenetic markings serve as a resource for RNA-mediated regulatory tools and additionally can represent impacts of environmental circumstances that may be heritable or not [Cuzin and Rassoulzadegan, 2010; Shapiro, 2014; Tognini *et al.*, 2015]. The role of epigenetics looks also like a memory tool which does not alter sequence structure but changes its regulation and function in multiple ways according environmentally induced adaptational needs [Mattick, 2010; Bredy *et al.*, 2011; Mercer and Mattick, 2013], and therefore it is also a main cause of diseases if regulatory networks get out of control [Marraffini and Sontheimer, 2010; Spadafora, 2015). Therefore it represents one of many kinds of natural genetic engineering for installing information via proteins and/or RNAs to DNA, i.e. converting the central dogma of molecular biology [Shapiro, 2009, 2011, 2014].

In the realm of artificial genetic engineering gene synthesis arose as a technique of synthetic biology for producing artificial genes in laboratories. In contrast to molecular cloning and polymerase chain reactions gene synthesis does not need pre-existing DNA, but is synthesised as double-stranded DNA without limit in terms of length or sequence content.

A more accessible technique is artificial genome editing, not solely genetic engineering. The assumption is that just as editing a written text in human language involves adding, removing, or replacing words in sentences, in genome editing the genome sequences are changed by adding, replacing, or removing nucleotides [Jasin, 1996; Lyons *et al.*, 2003]. For genome editing gene 'scissors' are used for deleting certain sequence structures/ genes to see what effect the 'knock out' of certain genes has. This goal is reached by site-specific endonucleases that are used as an appropriate tool for selective genome cleavage [Jasin, 1996; de Souza, 2012]. Such endonucleases make it possible to direct gene targeting. The three methods that are currently used are based on zinc finger endonucleases, TALEN- gen 'scissors' and the more recent CRIPRS/Cas9 technique that has been detected as an

adaptive immune system in prokaryotes in which small parts of the genome of natural genetic parasites are integrated into the host genome and serve to ward off similar genetic parasites. Artificial genome editing therefore uses identification and manipulation techniques that may have far-reaching and in extreme cases infinite consequences on germ cells of organisms that are manipulated accordingly [Iranzo *et al.*, 2013; Hsu *et al.*, 2013; Lin *et al.*, 2014; O'Connell *et al.*, 2014].

Here it should be said that artificial genome manipulation is confronted with a number of technical problems such as the low quality of oligonucleotides, faults in the syntax of sequences, and damage or problems in terms of nucleotide assembly. Additionally, overlapping regions may cause identity problems with newly synthesised genes and error correction methods could be optimised. Last but not least we have to rethink the possible consequences of manipulation techniques which were first mentioned in the Asilomar conference and are now the subject of current discourse [Berg *et al.,* 1975; Baltimore *et al.*, 2015; Sugarman, 2015] and which act on sequence syntax, not forgetting that natural genomes are the result of long-lasting selection processes in vivo, which means they happened within the context of an abundance of various lifeworlds together with rather different co-consortia, such as symbionts, bacteria and an abundance of viruses and infectious RNAs, all of them absent in in vitro technical set-ups [Villarreal, 2005; Ryan, 2009].

Therefore there is a crucial difference in the theoretical assumptions: is the genome the result of natural editing by competent agents that assemble a genetically conserved background resulting out of a rich evolutionary history of billions of years or is it solely the result of a variety of selection processes within some genetic drift passages of chance mutations in the realm of cell machineries that can be viewed as molecular bricks that can be restructured and rebuilt in a Lego-like fashion? In the first perspective we have a superficial nucleotide sequence grammar which can be epigenetically marked in different ways like a hidden, deep grammar that is not obvious in the superficial

grammar. In the second perspective there is only one superficial grammar, and it is visible, measurable and can be computed by algorithm-based procedures. The hidden deep grammar is not the focus here.

Because the consequences of the contradictory perspectives are subject to ethical debates, and the science of ethics is beyond the expertise of natural sciences, this is not further evaluated here.

3. The Old Concept: The Genetic Code in a Quantifiable Sequence Space

In the early 1960s it became increasingly clear that genetic information is stored in a molecular structure of nucleotide sequences termed the genetic code. It resembles all features of natural codes, an alphabet of nucleotides which can be assembled in only one reading direction and read, transcribed (into intermediate RNA) and translated at least into proteins which form organismal bodies, i.e. their parts and metabolism. The rules governing how gene alphabet characters are combined naturally (i.e. the molecular syntax) were identified by Erwin Chargaff (Chargaff rules) who demonstrated the results of his investigations to the young James Watson and Francis Crick. The latter afterwards detected the molecular structure of the double helix and Crick termed it a 'code without commas'. He also observed that information transfer direction is irreversible in the traditional 'central dogma of molcular biology': DNA — RNA — proteins — anything else [Crick, 1970].

At the same time information theory and cybernetic systems theory emerged and the genetic code was immediately interpreted in the light of these two emerging theories by molecular biologists and geneticists. Therefore the natural genetic code was viewed as a molecular structure that can be measured, explained and understood by natural laws, physics, chemistry and information theory [Schrödinger, 1944; Eigen, 1971]. Recently, Sydney Brenner argued that cells and living organisms represent the best examples of Turing and von Neumann machines [Brenner, 2012;

Witzany and Baluska, 2012]. However, the concept of Alan Turing and John von Neumann meanwhile became a 'Touring'machine, i.e. 'touring' through the history of science: no real Turing machine has been seen in reality since all the expected and visionary predicted beneficials of the last 70 years.

The reason why system theory and information theory are preferred in molecular biology and genetics is that there was a far-reaching discourse on building up an exact scientific language in contrast to non-scientific ones (metaphysics, vitalism) which would lead to exact science (natural sciences) in which only such sentences as were formalisable would fulfil the science criteria. Only formalisable sentences could depict material reality, i.e. reality built of physics and chemistry, and every entity of this reality would be formalisable in a mathematical 'Hilbert space'[Hilbert and Bernays, 1934, 1939; Whitehead and Russell, 1910, 1912, 1913] by unique coordinates that could be depicted in mathematical equations in principle. This concept later on was adapted to "biology as sequence space" [Eigen and Biebricher, 1988]. Built on these assumptions, systems theory and information theory were assumed to be the best methods for explaining the genetic information representing self-organised matter, i.e. the molecular structure of the genetic code [Eigen, 1971]. Both became privileged concepts for investigating the genetic code coherently, as shown by the importance of bioinformatics, biolinguistics, systems biology, mathematical biology, synthetic biology, i.e. quantifiable analyses of the features of the genetic code [Witzany, 2010].

4. Discredited Theoretical Assumptions in Molecular Biology

In the 1990s if not before the theoretical core assumptions in molecular biology and genetics changed dramatically. The central dogma of molecular biology that sequential information cannot be transferred from protein to either protein or nucleic acid was disproved in multiple examples [Shapiro, 2009, 2011]. This led to Crick's prediction that the wrong assumption ' ... would shake the

whole intellectual basis of molecular biology' [Crick, 1970]. Additionally, the one gene-one protein hypothesis was disproved. One gene can code for several proteins because the epigenetic marking of the gene sequence may cause several transcription and translation patterns. Of equal importance was the disproving of the assumption that gene sequences that do not code for proteins represent former evolutionary stages without any function, remaining as useless 'junk' DNA. However, we know that nearly all of the non-coding DNA is also transcribed into a variety of RNAs that are split up by several co-ordinated steps into small noncoding RNAs such as micro RNAs that fulfil a variety of essential functions in gene regulation [Mattick and Gagen, 2001; Mattick and Makunin, 2006; Mattick *et al.*, 2010].

Unexpectedly, the most powerful development was the comeback of virology. Although it was observed many years ago that '….life may have remained in the virus stage for many millions of years before a suitable assemblage of elementary units was brought together in the first cell' [Haldane, 1929], with the rise of molecular biology the idea re-emerged that viruses represent escaped parasites of cellular organisms that are non-essential parts of the tree of life. That viruses emerged earlier than cellular life was dismissed for decades. Empirical knowledge now indicates that several genomic features of viruses cannot be found in any cellular genome, which indicates an older evolutionary status [Villarreal, 2005; Koonin *et al.*, 2006; Villarreal and Witzany, 2010].

5. Essential Features of Natural Codes

After the aforementioned attempts to generate an exact scientific language to depict material reality by using formalisable equations to represent objective entities within a formalisable 'universe of entities' as proposed by Hilbert,Whitehead and Russell the theory of science discourse turned into pragmatics thanks to late Ludwig Wittgenstein. Wittgenstein demonstrated that the exact science language that he founded early in his famous 'Tractatus logico

philosophicus' was a fundamental error [Wittgenstein, 1953; Witzany, 2014a]. In contrast to artificial language constructions such as formalizable scientific languages natural everyday languages are the ultimate source material for investigations into how natural languages arise and function: consortia of competent living agents develop sign systems of various forms by themselves for cooperation and coordination of everyday behaviour, which means natural languages are inherently a kind of social interaction mediated by signs (indices, signals, symbols and behavioural embodiments that can express similar functions). Only once one cannot follow rules. Rule following is inherently a kind of customized social interactions. One biological entity alone could never emerge for the first time, with the consequence that the theoretical assumption of LUCA (last universal common ancestor) in terms of cellular life remains a chimera of false theoretical pre-assumptions [Villarreal and Witzany, 2010].

The semantics of signs, i.e. the meanings of the signs depend on the context in which signs are used by biological interacting groups [Witzany, 2010]. This means the same 'word' or alphabetic sequence can have multiple meanings within different contextual circumstances. The 'word' (or similar syntactic sign assemblies) has a visible superficial grammar, but the range of contextual usages may add several different meanings to the identical word grammar. This represents the deep grammar inherently interconnected with the situational context of the usage of a word that is not visible in the superficial grammar but can be used by living agents according to their different situational needs [Witzany and Baluska, 2012].

Additionally but of similar importance than this result of Wittgenstein, Kurt Gödel demonstrated in his incompleteness theorem that the assumptions Hilbert used to construct a contradiction-free axiomatic system in 'Hilbert space' are impossible in principle, because in natural language-using populations there is an inherent possibility of generating new sentences, new sign sequences that do not result from previous

ones and cannot be predicted by algorithm-based procedures [Gödel, 1931].

If the genetic code is not the result of replication errors in the self-organisation of matter, but an inherent active biological phenomenon, and additionally an ecosphere habitat for a rich lifeworld of RNA species that not only compete but cooperate, then we now need to focus on the new perspective of the natural genetic code [Mauricio, 2005; Brookfield, 2005; Le Rouzic *et al.*, 2007; Vennera *et al.*, 2009; Witzany, 2015a].

Most importantly, no natural language speaks itself; nor does a natural code code itself. There is an essential precondition for natural languages and codes, i.e. living agents which act as semiotic subjects; this means groups, societies, swarms that share the three levels of (syntactic, pragmatic and semantic) rules of language/code usage with which they organise and coordinate common behaviour. The relationship of living agents with their (historically evolved) real-life situation we term pragmatics [Witzany, 2014b]. Consortia of living agents share pragmatic rules to install sign-mediated interactions. It is important to note that semiotic rules — although quite conservative — may be changed by the user communities according to adaptational needs. This is the crucial difference of semiotic rules that determine sign usage to natural laws that cannot be changed but every entity underlies them in a strict sense [Witzany, 2015b].

In summary, living agents that cooperate and coordinate their behaviour via sign(al)s follow three levels of rules to combine signs correctly to generate more complex sign sequences (syntactic rules), choose behavioural patterns that are appropriate for fostering cooperative behaviour (pragmatic rules) and therefore determine the information content for the designation needs the signs serve (semantic rules).

Last but not least we have to look at how natural code users save energy costs. In natural codes we have a limited number of signs and a limited number of rules with which living agents generate and coordinate behaviour. Because natural language/code tools are limited, the information-bearing sequences may designate

several independent and even contradictory contents.. One word may have several different meanings, because living agents cannot invent new 'words' or sign sequences for every new situation or designation. Similar or equal combinations of signs, characters and words which result in sentences can be used as informational tools to transport different meanings. The phrase 'The shooting of the hunters', for example, cannot be understood unequivocably. In the one context this may indicate a common shooting of hunters of non-self targets, in the other it may mark dramatic misbehaviour [Witzany and Baluska, 2012]. The marking of syntactic sequences by marking tools is common practice in natural languages/codes and determines semantic content according to the needs of the pragmatic interacting agents.

To investigate syntactic sequences without knowing something about the real-life behaviour of code-using agents is senseless because syntactic structures do not represent unequivocably semantic meaning. Quantifiable analyses of signs, words or sequences cannot extract meaning. Only in a rather restricted quantifiable sense is this possible through sequence comparison with its known functions. All these features are absent in non-animate nature. If water freezes to ice no living agents or semiotic rules or signs are necessary and present.

6. Natural Genome Editing: What Does it Mean?

The genetic code in systems theory and information theory is not the result of interacting agents but of selection of replication errors (mutations) of biological macromolecules. Because Manfred Eigen assumes that information-bearing codes in macromoleclar systems, as well as in complex phenotypic systems such as human brains, represent self-organising matter, it is less difficult to move from a single macromolecule to a living cell than assume the transition of the single cell to an intelligent human being [Eigen, 1971, 2013]. Information in this sense is a molecular property within a dynamic theory of matter that gets its value through its self-reproductivity. Eigen's conclusion that there is no essential difference between

abiotic matter and biotic entities except the emergence of biological information by hypercycles of quasi-species is inevitable. Both depend on natural laws that govern physico-chemical cause and effects. A series of replication errors (mutations) of master copies leads to quasi-species, that are mutant distributions of primitive replicating entities. Such dynamic distributions of genomes that share genetic variation, competition and selection generate the fittest types (i.e. master copies) or in the extreme case error thresholds, i.e. excessively high mutation rates/variations, in that information cannot further reproduce in the case of excessive mutational load [Eigen, 1971].

Although the quasi-species concept predominated evolution biology for nearly half a century it is not an appropriate model for coherently explaining more recent empirical data on co-operative consortia of RNA groups and viruses or its "defectives" that co-operate [Villarreal and Witzany, 2013a,b). The evidence that the evolution, conservation and plasticity of genetic identities are the result of co-operative consortia of RNA stem loops being able to use natural code and edit this code, even with the generation of new sequences opens a new perspective on artificial genome editing as well [Witzany, 2011; Villarreal and Witzany, 2015].

Especially the ability to generate really new sequences (not out of previous ones) allows such groups constantly to infect other nucleic sequence-based agents, whether virus-like or cellular genomes. The generation of such new sequences by co-operating RNA stem loop groups lead to identity groups such as viruses that represent toxic codes and even counteracting antitoxins. Persistently (non-lytic) infected host organisms are the preferred habitat where former competing agent groups are unified in the basic behavioural motif of 'addiction modules' (*Gangen hypothesis*) that can be identified as TA, RM, ID modules; all of the former competing groups are now unified to form stable/ unstable modules that are counter-regulated and also provide immunity against related genetic parasites [Villarreal, 2009, 2011a, 2015]. In this way the result of unifying viruses and their defective parts (quasi-species consortia: qs-c) can explain the evolution,

conservation and plasticity of genetic identities more coherently than the previous quasi-species concept [Villarreal and Witzany, 2015].

7. Consequences for Artificial Genome Editing

Natural genome editing means changing nucleotide sequences actively, not to be colonised but to colonise, i.e. a passive chemical process (copying by complementary base pairing) vs. active editing of code (changing the nucleotide order to colonise/ manipulate former sequences). The groups of various counter-balanced ribozymes assemble single competences into a complex competence. It is important to note that previously deleted or fragmented RNA remnants may be re-used and re-integrated into group-building later [Villarreal and Witzany, 2013a, 2015, Villarreal, 2015].

Most importantly, this ensemble-building is context-dependent in terms of the history of the ecosphere: temperature (cold, hot), light (yes, no), water (fluid, icy), ph gradients, density, dry land, and combinations and intermediates may determine which ribozymatic features dominate, which are less dominant, which compete, which preclude each other and which cooperate. In particular, the intermediate stages in most cases cannot be defined in a formalisable way, which was a resulting problem in the above mentioned philosophy of science discourse [Witzany, 2007].

Additionally, the RNA group assemblies represent key features of ecological conditions. To 'survive', rapid changeability and less stability are necessary, because only permanent innovation of sequences guarantees the emergence of better colonisers or, in a cooperative way, the integration of parts of genetic parasites as useful weapons to defeat similar parasites through effective immune functions as represented by the various adaptive immune systems [Villarreal, 2011b; Moelling and Broecker, 2015]. In the light of natural genome editing error replication events (chance mutations) would not optimise but reduce the emergence of beneficial innovations. The other extreme, mutational overload

("error catastrophe") in natural genome editing, means: Too many innovations and sign-sequence generations cannot be shared in a community ecology of the RNA-based genome inhabitants because integration into the group competence within a real life world is too much new information tools. We know that from our experience: New words and terms may be usefull to understand new experiences or observations. The use of too many new terms confuses our competence to build society-based conventions regarding how to use such terms.

Now we have to see whether there are agents that are competent naturally to edit genetic code as sequence syntax and additionally mark the whole complex genome epigenetically. This must be coherent with molecular features, atomic structure, information processing and code editing rules, i.e. syntax, semantics and, most importantly, pragmatics, because the context determines semantics/meaning. So, what are these agents?

8. At the Core of Natural Genetic Novelty: Interacting RNA Consortia

Similarly interesting is this new perspective on the genome because it combines the atomic level and molecular level via nucleic acid sequences into a variety of unique and novel sequence combinations that are not object to algorithm-based procedures. The emergence of new genetic information is not the result of processes being subject to formalisable/mathematical equations, but the inherent feature of single-stranded RNA sequences which fold back and form stem-loop structures in a rather dynamic way, serving as a passive template or catalytically active agent, switching in between in unpredictable ways [Kumar and Joyce, 2003; Smit *et al.*, 2006; Gwiazda *et al.*, 2012; Müller *et al.*, 2012; Müller, 2015]. This is an important strategy for unlimited progression of the interplay between infectious agents, host organisms that conserve this by integrating genetic information of the identity of the infectious agents to ward off related infectious agents, and the generation of new genetic information which again

may "overrule" such immune functions [Villarreal, 2009; Villarreal, 2011a,b].

The result is a completely new level of information content that is absent in inanimate nature. In contrast to agglomeration in pure chemical nature we have a finite number of characters of the nucleotide alphabet within an infinite combinatorial space of nucleic acid sequences. On the single-stranded RNA level we can see the formation of loops that fold back to the single strand, to build a double stranded (base-paired) RNA with a single-stranded loop.

In this context the RNA loop can generate an identity which other based/non-base-paired assemblies do not share. Additionally, the RNA stem loop has a part, the non-pairing loop, with a rather "sticky" section and can interact with other RNAs of the same or other RNA groups with similar non-base pairing loops or even single-stranded remants of former RNA agents [Witzany, 2014c]. They can be found by testing other sequences, identifying them as appropriate binding sequences or rejecting them because sequence structure does not fit. This interaction motif can be termed RNA sensing or RNA monitoring action. Here we can find some sort of identification competence where the single RNA stem loop or a group or RNA stem loops represents a kind of biological 'self' which can identify other 'self' or 'non-self' RNA groups to cooperate or reject [Villarreal, 2009a,b, 2011a]. This is the reason why RNA groups may act in an active catalytic way or simply be a passive template for replication. Interestingly, thanks to this inherent double function they may change both functions in a rather non-predictive way. Whereas DNA forms a predictable double helix, RNA comprises single strands that fold up into loops, bulges, pseudo-knots, hammerheads, hairpins and other motifs. These structures flip and twist between different forms in a non-predictive manner.

With the identification of the essential agents of natural genetic engineering and natural genome editing, RNA consortia of various groups, and their inherent ability to build base-pairing parts and non-base pairing but sensing and monitoring loops we

have identified the core agents of genetic novelty, i.e. of evolutionary processes. It is important to note that this is not error replication (mutation) but represents real evolutionary history-derived and stored competence to generate new sequence motifs.

Now, after identifying the agents that naturally edit genetic code and are the main source of genetic novelty let us have a look at the core process of genetic novelty. It is not just an error replication event (mutation) as suggested by former theoretical concepts which view this as a process of self-organisation of matter. If this is a process which depends on interactions of RNA groups then it is inherently interwoven with self/non-self interactions, i.e. group building, integration of appropriate beneficial agents into groups or preclusion, deletion of RNAs or RNA groups which do not fit into the pre-existing group [Vaidya *et al.*, 2012]. This means the process of genetic novelty is interwoven into a more complex process of essential group identity, preserving group identity and attack against group identity infection, destruction or damage [Osborn and Boltner, 2002; Huda *et al.*, 2010; Villarreal, 2012]. On the other hand, we then have to look at the various motifs and techniques that are available in RNA groups to prevent infection events such as the generation of a diversity of immunity tools or weapons to attack and destroy invading omnipresent genetic parasites (such as endonucleases) [Villarreal, 2012; Moelling, 2013]. Additionally there must be tools to integrate beneficial group members to build more complex groups. This is the most powerful ligase tool.

More generally, the crucial difference between biological identities and non-biological identities is that the first are based on a biological code by agents that share code-using rules, whereas the latter miss both, i.e. no biological codes or competent code-using agent groups are present.

First, we should look at the more basic process that is loop building within RNA consortia. This happens when within a stem or a loop the base-pairing nucleotides are broken up into a section of non-base-pairing (single-stranded regions of) nucleotides. The results are various. They may reach from plus strand variation to

negative strand variation, both influencing all RNA consortia/ group interactions and e.g., in the editosome or in the spliceosome a complete loss of function of the former entire agents [Villarreal and Witzany, 2013a,b; Witzany, 2011]. The ongoing generation of such loops is the natural core competence for producing novelty, for which no immune function exists and which serves as an evolutionary tool to invade or, as a persistent integrated feature in a host, preclude infectious agents.

The various consequences in a group of RNA agents such as the ribosomal subunits, the editosome, the spliceosome and others with multiple RNA stem loops that build a cooperative agent are therefore algorithmically unpredictable, because of the unlimited possibilities of combinations of certain group identities which may result from a single broken stem which then builds a loop. If we look for example at the ribosomal subunits the folding (pragmatics) of the sub-groups determines their functionality, not their sequence syntax [Bokov and Steinberg, 2009; Harish and Caetano-Anolles, 2012; Petrov *et al.*, 2013].

9. Comeback of the Century? Viruses and Virus-Like RNA Agents Interact as Cooperative Groups

Current research results additionally indicate that viruses are the most abundant biological agents on this planet (10 times more abundant than cellular genomes), and only viruses assemble all known features of the genetic code, such as double-stranded or single-stranded RNA or DNA (+ and – stranded) [Forterre and Prangishvili, 2009; Geuking *et al.*, 2009; Rossinck, 2011, 2012; Jalasvuori, 2012; Koonin and Dolja, 2014; Koonin *et al.*, 2015]. In prokaryotes phages are nearly omnipresent and massively determine their host gene word order. Also, the eukaryotic nucleus resembles a variety of large dsDNA virus features. In every mitochondria, endoplasmic reticulum or even plasmids we can find persistent viral parts. The endogenous retroviruses (active and/or defective) play crucial roles in the evolution of higher

animals [Canchaya *et al.*, 2004; Briones *et al.*, 2006; Carbonell *et al.*, 2012; Crespi and Nosil, 2013].

We can identify virally derived insertions that remain as defectives such as non-coding RNAs essential in gene regulation as intronic regions that are spliced out during exon assembly.

Some persistent viruses/virus-derived parts which are counter-regulated by opposing or former competing genetic parasites have been identified such as DNA viruses, DNA transposons, RNA viruses, non-retroviral RNA viruses, endogenous retroviruses, LTRs retrotransposons, non-LTRs (SINEs, LINEs, ALUs), group II introns, group I introns. All of these active parts play essential roles in natural nucleotide recombination techniques such as those used in DNA/RNA structuring and restructuring, amplifying or silencing functions, sub-steps of transcription (post-transcriptional RNA editing, RNA splicing, ribosome assembly), translation, DNA replication, chromatin organisation, epigenetic markings and modifications, DNA repair [Xiong and Eickbush, 1988; Baranowski *et al.*, 2001; Sun *et al.*, 2006; Weiner, 2006; Feschotte, 2008; Perot *et al.*, 2012; Cowley and Oakey, 2013; Swart and Nowacki, 2015, Zimmerly and Semper, 2015]. Perhaps the best examples of persistent life style of viruses are represented by the organisation of the various forms of adaptive and innate immunity systems such as CRISPRs/Cas or the amaizingly complex VDJ immune system [Villarreal, 2009 a,b, 2011b].

All these examples show that the genome is not merely a molecular structure with a storage function but rather an ecosphere habitat with an abundance of RNA-derived settlers that compete for a rather limited resource [Witzany, 2012]. Most interestingly, to get access to this limited resource some cooperative behavioural patterns have been selected whereby formerly competing agents find a way to cooperate, to counter-regulate within the host genome. This new identity co-oparation of former competing genetic parasites also may lead to new identities of host tissues, organs or organisms, really evolutionary novelty. It is possible to imagine how different tissues evolved in quite different species;

this is a coherent event because of an abundance of persistent (non-lytic) viruses which share a tissue specificity not a species specificity [Villarreal, 2005]. In single infection events up to 100 new genes can transfer to a new host. This is not a small step as in error replication events (chance mutations) but an evolutionary non-random drive with far-reaching consequences as documented by e.g. the retroviral infections that lead to the transfer of syncytin genes, which resulted into the evolutionary novelty of placental mammals [Villarreal, 2005; Perot *et al.*, 2012].

Unexpectedly, the controversial theoretical concepts of evolutionary novelty being essential for diversity and its selection processes are no longer the undirected or directed mutation narrative, nor teleological vitalism metaphysics (more recently 'intelligent design'), nor the molecular biological self-organisation of matter (Eigen-Schuster narrative), nor the increasing complexity of a self-emerging property of systems (Kauffman narrative), but natural genetic content organisation by competent microbial/viral agents that cooperate for their survival goals which may coincide with those of their hosts as documented in the variety of endosymbiotic evolutionary processes [Witzany, 2006].

10. Conclusions

The old success story of genetic engineering and the more recent dawn of genome editing faces some technical problems. On the other hand, the current debate on the ethical justification of these techniques of genome manipulation is still open. Of similar importance are the dramatic changes in theoretical pre-assumptions together with recent empirical knowledge about the capabilities of RNA consortia, persistent viruses and other infectious genetic parasites. The old narrative of molecular entities that assemble according to physico-chemical properties of matter dictated by natural laws such as thermodynamics, quantum physics and chemical binding is increasingly enriched by the finding that groups of RNA stem loops generate an abundance of nucleic acid code-based consortial interactions. We find single RNA stem loops

with fast-changing identities that build groups such as ribosomal subunits, editosome, spliceosome as active (catalytic) or passive (template-like) agents switching in a non-predictable way. With these behavioural motifs the emergence of biological identity (self/non-self identification competence) occurs. RNA groups are able to act as de novo producers of nucleic acid sequences, identify sequence-specific target sites, coherently integrate such sequences into pre-existing ones (without destruction of former content arrangements), recombine according to adaptational needs and mark sequence sites to vary meaning epigentically or identify sequences to be marked for excision or deletion. In all these processes the genetic identity of the genetic parasite and/or the host genome may vary, with far-reaching consequences in terms of the function, co-operation and coordination of various regulatory networks. Natural genome editing is therefore far from being a random-like process as a result of error replication (mutations). Artificial genome editing will have to integrate the agent-based perspective into its theoretical assumptions as well as the contextual real lifeworlds of these agents to achieve a more realistic and integrative view on the empirical data currently available. The perspective on natural genetic information processing is changing dramatically.

References

Baltimore, D. *et al.* (2015). A prudent path forward for genomic engineering and germline gene modification. *Science* 348, pp. 36-38.

Baranowski, E., Ruiz-Jarabo, C.M and Domingo, E. (2001). Evolution of cell recognition by viruses. *Science* 292, pp. 1102-1105.

Barlow, D.P. (2011). Genomic imprinting: a mammalian epigenetic discovery model. *Annu. Rev. Genet.* 45, pp. 379-403.

Berg, P., Baltimore, D., Brenner, S., Roblin, R.O. and Singer, M.F. (1975). Summary statement of the Asilomar conference on recombinant DNA molecules. *PNAS* 72, pp. 1981-1984.

Bokov, K. and Steinberg, S.V. (2009). A hierarchical model for evolution of 23S ribosomal RNA. *Nature* 457, pp. 977-980.

Bredy, T.W., Lin, Q., Wie, W., Baker Andresen. D. and Mattick, J. (2011). Micro RNA regulation of neural plasticity and memory. *Neurobiol. Learn. Mem.* 96, pp. 89-94.
Brenner, S. (2012). Life's code script. *Nature* 482, p. 461.
Briones, C., de Vicente, A., Molina-París, C. and Domingo, E. (2006). Minority memory genomes can influence the evolution of HIV-1 quasispecies in vivo. *Gene* 384, pp. 129-138.
Brookfield, J.F.Y. (2005). The ecology of the genome. Mobile DNA elements and their hosts. *Nat. Rev. Genet.* 6, pp. 128-136.
Canchay, C., Fournous, G. and Brussow, H. (2004). The impact of prophages on bacterial chromosomes. *Mol. Microbiol.* 53, pp. 9–18.
Carbonell, A., Flores, R. and Gago, S. (2012). *Hammerhead Ribozymes Against Virus and Viroid RNAs*. Eds. Erdmann, V.A. and Barciszewski, J., "From nucleic acids sequences to molecular medicine", Springer, Berlin, Heidelberg, pp. 411–427.
Cech, T.R. and Steitz, J.A. (2014). The noncoding RNA revolution — trashing old rules to forge new ones. *Cell* 157, pp. 77–94.
Cowley, M. and Oakey, R.J. (2013). Transposable elements re-wire and fine-tune the transcriptome. *PLoS Genet.* 9, e1003234.
Crespi, B. and Nosil, P. (2013). Conflictual speciation: species formation via genomic conflict. *Trends Ecol. Evol.* 28, pp. 48–57.
Crick, F. (1970). Central dogma of molecular biology. *Nature* 227, pp. 561-563.
Curcio, M.J. and Derbyshire, K.M. (2003). The outs and ins of transposition: from mu to kangaroo. *Nat. Rev. Mol. Cell Biol.* 4, pp. 856-877.
Cuzin, F. and Rassoulzadegan, M. (2010). Non-Mendelian epigenetic heredity: gametic RNAs as epigenetic regulators and transgenerational signals. *Essays Biochem.* 48, pp. 101-106.
de Souza, N. (2012). Primer: genome editing with engineered nucleases. *Nat. Meth.* 9, p. 27.
Eigen, M. (1971). Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* 58, pp. 465-523.
Eigen, M. (2013). *From Strange Simplicity to Complex Familiarity: A Treatise on Matter, Information, Life and Thought*. Oxford University Press, Oxford, UK.
Eigen, M. and Biebricher, C. K. (1988). *Sequence Space and Quasipecies Distribution*.
eds. Domingo, E., Ahlquist, P. and and Holland, J.J. "RNA genetics", vol. 3. CRC Press, Boca Raton, USA, pp. 211-245.
Feschotte, C. (2008). Transposable elements and the evolution of regulatory networks. *Nat. Rev. Genet*. 9, pp. 397-405.
Forterre, P. and Prangishvili, D. (2009). The great billion-year war between ribosome- and capsid-encoding organisms (cells and viruses) as the major source of evolutionary novelties. *Ann. N.Y. Acad. Sci.* 1178, pp. 65-77.
Geuking, M.B. *et al.* (2009). Recombination of retrotransposon and exogenous RNA virus rersults in nonretroviral cDNA integration. *Science* 323, pp. 393-396.
Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme. *Monatsh. Math. Phys.* 38, pp. 173-198.
Gwiazda, S., Salomon, K., Appel, B. and Mueller, S. (2012). RNA self-ligation: from oligonucleotides to full length ribozymes. *Biochimie* 94, pp. 1457-1463.
Haldane, J.B.S. (1929). The origin of life. *Rationalist Ann.* 148, pp. 3-10.

Harish, A. and Caetano-Anolles, G. (2012). Ribosomal history reveals origins of modern protein synthesis. *PLoS One* 7, e32776.

Hilbert, D. and Bernays, P. (1934/1939). *Grundlagen der Mathematik*. Vol.1/2. Springer, Berlin/New York.

Hsu, P.D. *et al.* (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotech.* 31, pp. 827-832.

Huda, A., Mariño-Ramírez, L. and Jordan, I.K. (2010). Epigenetic histone modifications of human transposable elements: genome defense versus exaptation. *Mob. DNA* 25, p. 2.

Iranzo, J. *et al.* (2013). Evolutionary dynamics of the prokaryotic adaptive immunity system CRISPR-Cas in an explicit ecological context. *J. Bacteriol.* 195, pp. 3834-3844.

Jalasvuori, M. (2012). *Revolutionary Struggle for Existence: Introduction to Four Intriguing Puzzles in Virus Research*, ed. Witzany, G. Viruses: Essential Agents of Life", Springer, Dordrecht, NL, pp. 1-19.

Jasin, M. (1996). Genetic manipulation of genomes with rare-cutting endonucleases. *Trends Genet.* 12, pp. 224-228.

Jirtle, R.L. (2009). Epigenome: the program for human health and disease. *Epigenomics* 1, pp. 13-16.

Koonin, E.V. (2009). On the origin of cells and viruses: primordial virus world scenario. *Ann. N.Y. Acad. Sci.* 1178, pp. 47-64.

Koonin, E.V., Senkevich, T.G. and Dolja, V.V. (2006). The ancient Virus World and evolution of cells. *Biol. Direct*. 1, p. 29.

Koonin, E,V, and Dolja, V.V. (2014). Virus world as an evolutionary network of viruses and capsidless selfish elements. *Microbiol. Mol. Biol. Rev.* 78, pp. 278-303.

Koonin, E., Dolja, V.V. and Krupovic, M. (2015). Origins and evolution of viruses of eukaryotes: The ultimate modularity. *Virology* 479/480, pp. 2-25.

Kumar, R.M. and Joyce, G.F. (2003). A modular, bifunctional RNA that integrates itself into a target RNA. *PNAS* 100, pp. 9738-9743.

Le Rouzic, A., Dupas, S. and Capy, P. (2007). Genome ecosystem and transposable elements species. *Gene* 390, pp. 214-220.

Lin, S., Staahl, B.T., All, R.K. and Doudna, J.A. (2014.) Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife* 3, e04766.

Lyons, A.J. and Robertson, H.D. (2003). Detection of tRNA-like structure through RNase P cleavage of viral internal ribosome entry site RNAs near the AUG start triplet. *J. Biol. Chem.* 278, pp. 26844-26850.

Marraffini, L.A. and Sontheimer, E.J. (2010). Self versus non-self discrimination during CRIPR RNA-directed immunity. *Nature* 463, pp. 568-571.

Mattick, J.S. (2009). The genetic signatures of non-coding RNAs. *PLoS Genet.* 5, e1000459.

Mattick, J.S. (2010). RNA as a 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, pp. 548-552.

Mattick, J.S. and Gagen, M.J. (2001). The evolution of controlled multitasked gene networks: the role of introns and other noncoding RNAs in the development of complex organisms. *Mol. Biol. Evol.* 18, pp. 1611-1630.

Mattick, J.S. and Makunin, I.V. (2006). Non-coding RNA. *Hum. Mol. Genet.* 1, pp. 17-29.

Mattick, J.S., Taft, R.J. and Faulkner, G.J. (2010). A global view of genomic information-moving beyond the gene and the master regulator. *Trends Genet.* 26, pp. 21-28.

Mauricio, R. (2005). Can ecology help genomics: The genome as ecosystem. *Genetica* 123, pp. 205-209.

Mercer, T.R. and Mattick, J.S. (2013). Structure and function of long non- coding RNAs in epigenetic regulation. *Nat. Struct. Mol. Biol.* 20, pp. 300-307.

Moelling, K. (2013). What contemporary viruses tell us about evolution: a personal view. *Arch. Virol.* 158, pp 1833-1848.

Moelling, K. and Broecker, F. (2015). The reverse transcriptase-RNase H: from viruses to antiviral defense. *Ann. N.Y. Acad. Sci.* 1341, pp. 126-135.

Mruk, I. and Kobayashi, I. (2014). To be or not to be: regulation of restriction-modification systems and other toxin–antitoxin systems. *Nucleic Acids Res.* 42, pp. 70-86.

Mueller, S. (2015). Engineering of ribozymes with useful activities in the ancient RNA world. *Ann. N.Y. Acad. Sci.* 1341, pp. 54-60.

Mueller, S., Appel, B., Krellenberg, T., Petkovic, S. (2012). The many faces of the hairpin ribozyme: Structural and functional variants of a small catalytic RNA. *IUBMB. Life* 64, pp. 36-47.

O'Connell, M.A., Oakes, B.J., Sternberg, S.H., East-Seletsky, A., Kaplan, M. and Doudna, J.A. (2014). Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature* 516, pp. 263-266.

Osborn, A.M. and Boltner, D. (2002). When phage, plasmids, and transposons collide: genomic islands, and conjugative — and mobilizable — transposons as a mosaic continuum. *Plasmid* 8, pp. 2002-2012.

Perot, P., Bolze, P.A. and Mallet, F. (2012). *From viruses to genes: syncytins*, ed. Witzany, G. "Viruses: Essential Agents of Life", Springer, Dordrecht, NL, pp. 325-361.

Petrov, A.S., Bernier, C.R., Hershkovits, E., Xue, Y., Waterbury, C.C. *et al.* (2013). Secondary structure and domain architecture of the 23S and 5S rRNAs. *Nucleic Acids Res*. 41, pp. 7522-7535.

Roossinck, M.J. (2011). The good viruses: viral mutualistic symbioses. *Nat. Rev. Microbiol.* 9, pp. 99-108.

Roossinck, M.J. (2012). *Persistent Plant Viruses: Molecular Hitchhikers or Epigenetic Elements?* ed. Witzany, G. "Viruses: Essential Agents of Life", Springer, Dordrecht, NL, pp. 177-186.

Ryan, F. (2009). *Virolution*. Harper Collins, London, UK.

Schrödinger, E. (1944). *What is Life*? *The Physical Aspect of the Living Cell*. Cambridge University Press, London, UK.

Shapiro, J.A. (2009). Revisting the central dogma in the 21st century. *Ann. N.Y. Acad. Sci*. 1178, pp. 6-28.

Shapiro, J.A. (2011). *Evolution: A View from the 21st Century*. FT Press, Washington DC, USA.

Shapiro, J.A. (2014). Epigenetic control of mobile DNA as an interface between experience and genome change. *Frontiers Genet.* 5, pp. 1-16.
Slotkin, R.K. and Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.* 8, pp. 272-285.
Smit, S., Yarus, M. and Knight, R. (2006). Natural selection is not required to explain universal compositional patterns in rRNA secondary structure categories. *RNA* 12, pp. 1-14.
Spadafora, C. (2015). A line-1-encoded reverse transcriptase-dependent regulatory mechanism is active in embryogenesis and tumorigenesis. *Ann. N.Y. Acad. Sci*. 1341, pp. 164-171.
Swart, E.C. and Nowacki, M. (2015). The eukaryotic way to defend and edit genomes by sRNA-targeted DNA deletion. *Ann. N.Y. Acad. Sci.* 1341, pp. 106-114.
Sugarman, J. (2015). Ethics and germline gene editing. *EMBO Rep.* 16, pp. 879-880.
Sun, F.J., Fleurdepine, S., Bousquet-Antonelli, C., Caetano-Anolles, G. and Deragon, J.M. (2006). Common evolutionary trends for SINE RNA structures. *Trends Genet.* 23, pp. 26-33.
Tang, W.W. *et al.* (2015). A unique gene regulatory network resets the human germline epigenome for development. *Cell* 161, pp. 1453-1467.
Tognini, P. *et al*. (2015). Experience-dependent DNA methylation regulates plasticity in the developing visual cortex. *Nat. Neurosci.* 18, pp. 956-958.
Vaidya, N., Manapat, M.L., Chen, I.A., Xulvi-Brunet, R., Hayden, E.J. and Lehman, N. (2012). Spontaneous network formation among cooperative RNA replicators. *Nature* 49, pp. 72-77.
Vennera, S., Feschotte, C. and Biemonta, C. (2009). Transposable elements dynamics: toward a community ecology of the genome. *Trends Genet*. 25, pp. 317-323.
Villarreal, L.P. (2005). *Viruses and the Evolution of Life*. ASM Press, Washington, USA.
Villarreal, L.P. (2009a). *Origin of Group Identity: Viruses, Addiction and Cooperation*. Springer, New York.
Villarreal, L.P. (2009b). The source of self: genetic parasites and the origin of adaptive immunity. *Ann. N.Y. Acad. Sci.* 1178, pp. 194-232.
Villarreal, L.P. (2011a). *Viruses and Host Evolution: Virus-Mediated Self-Identity*. ed. Lopez-Larrea, C. "Self and Non-Self", Landes Bioscience and Springer Science-Business Media, Austin, USA, pp. 185-217.
Villarreal, L.P. (2011b). Viral ancestors of antiviral systems. *Viruses* 3, pp. 1933–1958.
Villarreal, L.P. (2012). *The Addiction Module as a Social Force*. ed. Witzany, G. "Viruses: Essential Agents of Life", Springer, Dordrecht, NL, pp. 107-145.
Villarreal, L.P. (2015). Force for ancient and recent life: viral and stem-loop RNA consortia promote life. *Ann. N.Y. Acad. Sci.* 1341, pp. 25-34.
Villarreal, L.P. and Witzany, G. (2010). Viruses are essential agents within the roots and stem of the tree of life. *J. Theor. Biol.* 262, pp. 698-710.
Villarrreal, L.P. and Witzany, G. (2013a). The DNA habitat and its RNA inhabitants. At the dawn of RNA sociology. *Genom. Ins.* 6, pp. 1-12.
Villarreal, L.P. and Witzany, G. (2013b). Rethinking quasispecies theory: From fittest type to cooperative consortia. *World J. Biol. Chem.* 4, pp. 79-90.
Villarreal, L.P. and Witzany, G. (2015). When competing viruses unify: evolution, conservation, and plasticity of genetic identities. *J. Mol. Evol.* 80, pp. 305-318.

Weiner, A.M. (2006). *SINEs and LINEs: Troublemakers, Saboteurs, Benefactors, Ancestors*. Eds. Gesteland, R.F., Cech, T.R. and Atkins, J.F. "The RNA World", 3rd edn. Cold Spring Harbor Laboratory Press, New York, USA, pp. 507-534.

Werner, A., Piatek, M.J. and Mattick, J.S. (2015). Transpositional shuffling and quality control in male germ cells to enhance evolution of complex organisms. *Ann. N.Y. Acad. Sci.* 1341, pp. 156-163.

Whitehead, A.N. and Russell, B. (1910,1912,1913). *Principia Mathematica*, 3 vols., Cambridge University Press, Cambridge, UK.

Wittgenstein, L. (1953). *Philosophical Investigations*. Basil Blackwell, Oxford, UK.

Witzany, G. (2006). Serial Endosymbiotic Theory (set): the biosemiotic update. *Acta Biotheor*. 54, pp. 103-117.

Witzany, G. (2007). *From Biosphere to Semiosphere to Social Lifeworlds. Biology as an Understanding Social Science.* Ed. Barbieri, M. "Biosemiotic research trends", Nova Science Publishers, New York, USA, pp. 185–213.

Witzany, G. (2009). Noncoding RNAs: Persistent viral agents as modular tools for cellular needs. *Ann. N.Y. Acad. Sci.* 1178, pp. 244-267.

Witzany, G. (2010). *Biocommunication and Natural Genome Editing*. Springer, Dordrecht, NL.

Witzany, G. (2011). The agents of natural genome editing. *J. Mol. Cell Biol*. 3, pp. 181–189.

Witzany, G. (2012). *From Molecular Entities to Competent Agents: Viral Infection-derived Consortia act as Natural Genetic Engineers,* Ed. Witzany, G. "Viruses: Essential agents of life", Springer, Dordrecht, NL, pp 407-419.

Witzany, G. (2014a). *Language and Communication as Universal Requirements for Life*. Ed. Kolb, V. "Astrobiology. An evolutionary approach", CRC Press, Boca Raton, USA, pp. 407-419.

Witzany, G. (2014b). Pragmatic turn in biology: from biological molecules to genetic content operators. *World J. Biol. Chem.* 5, pp. 279-285.

Witzany, G. (2014c). RNA sociology: Group behavioral motifs of RNA consortia. *Life* 4, pp. 800-818.

Witzany, G. (ed.) (2015a). *The DNA Habitats and Their RNA Inhabitants*. John Wiley & Sons, Hanover, USA.

Witzany, G. (2015b). Life is physics and chemistry and communication. *Ann. N.Y. Acad. Sci.* 1341, pp. 1-9.

Witzany, G. and Baluska, F. (2012). Life's code script does not code itself. The machine metaphor for living organisms is outdated. *EMBO Rep.* 13, pp. 1054-1056.

Wolstenholme, G. (ed.) (1963). *Man and his future. A Ciba Foundation Volume*. J. & A. Churchill Ltd. London, UK.

Xiong, Y. and Eickbush, T.H. (1988). Similarity of reverse transcriptase-like sequences of viruses, transposable elements and mitochondrial introns. *Mol. Biol. Evol.* 5, pp. 675-690.

Zimmerly, S. and Semper, C. (2015). Evolution of group II introns. *Mob. DNA* 6, p. 7.