

UC Irvine

UC Irvine Previously Published Works

Title

Rethinking quasispecies theory: From fittest type to cooperative consortia

Permalink

<https://escholarship.org/uc/item/2b32v3hh>

Journal

World Journal of Biological Chemistry, 4(4)

ISSN

1949-8454

Authors

Villarreal, Luis P
Witzany, Guenther

Publication Date

2013

DOI

10.4331/wjbc.v4.i4.79

Peer reviewed

Rethinking quasispecies theory: From fittest type to cooperative consortia

Luis P Villarreal, Guenther Witzany

Luis P Villarreal, Department of Molecular Biology and Biochemistry, University of California, Irvine, CA 92697, United States

Guenther Witzany, Telos-Philosophische Praxis, Buermoos 5111, Austria

Author contributions: All authors contributed equally to this work.

Correspondence to: Dr. Guenther Witzany, Telos-Philosophische Praxis, Vogelsangstrasse 18c, Buermoos 5111, Austria. witzany@sbg.at

Telephone: +43-6274-6805 Fax: +43-6274-6805

Received: June 29, 2013 Revised: August 20, 2013

Accepted: September 3, 2013

Published online:

Abstract

Recent investigations surprisingly indicate that single RNA "stem-loops" operate solely by chemical laws that act without selective forces, and in contrast, self-ligated consortia of RNA stem-loops operates by biological selection. To understand consortial RNA selection, the concept of single quasi-species and its mutant spectra as drivers of RNA variation and evolution is rethought here. Instead, we evaluate the current RNA world scenario in which consortia of cooperating RNA stem-loops (not individuals) are the basic players. We thus redefine quasispecies as RNA quasispecies consortia (qs-c) and argue that it has essential behavioral motifs that are relevant to the inherent variation, evolution and diversity in biology. We propose that qs-c is an especially innovative force. We apply qs-c thinking to RNA stem-loops and evaluate how it yields altered bulges and loops in the stem-loop regions, not as errors, but as a natural capability to generate diversity. This basic competence - not error - opens a variety of combinatorial possibilities which may alter and create new biological interactions, identities and newly emerged self identity (immunity) functions. Thus RNA stem-loops typically operate as cooperative modules, like members of so-

cial groups. From such quasispecies consortia of stem-loop groups we can trace a variety of RNA secondary structures such as ribozymes, viroids, viruses, mobile genetic elements as abundant infection derived agents that provide the stem-loop societies of small and long non-coding RNAs.

© 2013 Baishideng. All rights reserved.

Key words: Quasispecies; RNA stem-loops; Infectious agents; Cooperative interactions; Evolution

Core tip: Single RNA stem-loops operate solely by chemical laws that act without selective forces, and in contrast, self-ligated consortia of RNA stem-loops operates by biological selection. To understand consortial RNA selection, the concept of single quasi-species and its mutant spectra as drivers of RNA variation and evolution is rethought here. Instead, we evaluate the current RNA world scenario in which consortia of cooperating RNA stem-loops (not individuals) are the basic players. We thus redefine quasispecies as RNA quasispecies consortia (qs-c) and argue that it has essential behavioral motifs that are relevant to the inherent variation, evolution and diversity in biology.

Villarreal LP, Witzany G. Rethinking quasispecies theory: From fittest type to cooperative consortia. *World J Biol Chem* 2013; 4(4): 000-000 Available from: URL: <http://www.wjgnet.com/1949-8454/full/v4/i4/00.htm> DOI: <http://dx.doi.org/10.4331/wjbc.v4.i4.00>

INTRODUCTION

Our mainstream picture of RNA virus evolution is determined by the quasispecies concept of Manfred Eigen. It was a very helpful theoretical model within the framework of physical chemistry to paint a picture of

viruses with an evolution rate a multitude faster than evolution of cellular organisms. The core assumptions of quasispecies concept^[1-3] are rather clear: (1) there is little knowledge about the origin of viruses or their role in the evolution of the biosphere. Viruses are selfish genetic elements that likely originated out of host organisms (escape theory); (2) viruses represent molecules that generate “information” through mutagenesis, *i.e.*, replication-errors; (3) equations of mass action law govern chemical equilibrium in closed chemical systems. Percentage of components within this is determined by these equations; (4) in an environment of high mutation rate and thermodynamic conditions far from equilibrium, self-replicating entities reach maximum reproductive fitness; (5) the self replication entity is not a single molecule but a “cloud” that consists of variant reproductive molecules; (6) the distribution of these “clouds” in systems far from equilibrium depends on master fittest type and mutant spectra, *i.e.*, dominant replicators, mutants closely related, and not closely related. Each of its sequence-syntax occupies a unique position in the sequence space; and (7) because the genetic code is a real language - not just a metaphor - information- and system-theoretical assumptions based on mathematical concepts of language therefore are the appropriate tools to explain quasispecies evolutionary dynamics.

If we look at the current empirical data, these assumptions do not explain many interactional behaviors and broad influences that have been observed in viral studies. In addition, the following issues are relevant: (1) viruses play large and essential roles in the biosphere evolution of host^[4-7]. Phylogenetic analyses and comparative genomics suggest that viruses are older than cellular life^[8-10]; (2) most viruses are symbiotic or even endosymbiotic inhabitants of cellular host organisms that can promote adaptability for the host^[11-15]; (3) information representing natural codes is not the result of errors in replication processes. All empirical data indicate that natural codes depend on consortia of living agents that generate and use codes for information storage and interactional signalling. Natural codes are essential prerequisites of everyday social interactions. If errors in natural code use occur, this does not enrich information but nearly all empirical data indicate information damage, deformation and decreased informational content. Today it seems rather curious to ground evolution of complexity on error and damage^[16-19]; and (4) coherent with Ludwig Wittgensteins assumptions, the meaning of natural code sequences is its use (functional activity), *i.e.*, depends on the context, not on its syntax^[20]. Accordingly the meaning (functional activity) of syntactically identical viral nucleic acid sequences varies according to context such as epigenetic imprinting^[21-23].

These observations have serious consequences for the “every variant” concept of Eigen’s quasi-species that include: (1) biological information of nucleic acid sequences does not occupy a unique position in sequence space but depends on contextual use; (2) because of its con-

text-dependency biological information cannot be sufficiently described by information theory or similar mathematically based concepts of language; (3) (Evolutionary) algorithm-based machines cannot provide contextual real life simulations; (4) sequence space of real life nucleotide sequences is not the result of random assemblies; and (5) viral cloud building in natural habitats occurs different in comparison to abiotic molecule assembly.

Additionally quasispecies that cooperate can have various behavioural motifs not present in a pure physical chemistry. They can (1) compete with and exclude related populations; (2) have minority populations that are crucial for overall fitness^[24,25]; (3) can display heterogeneity important for fitness that is not observed in the consensus type^[26]; (4) can suppress their own replication through lethal defection^[27]; and (5) can be composed of members that can complement and interfere with replication of the collective and many of these features can be observed in clinical infections such as humans with hepatitis C virus^[28].

Therefore if quasispecies evolve with the above characteristics and are thus different from prior mathematic models, what then is an appropriate description that is also coherent with the abiotic/biotic split in animated nature? If species is the appropriate term in biology to describe essential common features of related groups of cellular organisms, then quasispecies remains an appropriate term to describe related groups of subcellular agents that play essential roles in evolution of the biosphere.

To denote the above crucial differences to the 20th century concept with its paradigmatic core of master copy (fittest type) and mutant spectra (variants) we propose the concept of quasispecies consortia (qs-c) in which behavioural motifs of cooperative RNA-agents are at the foundation of basic capabilities generating sequence space of de novo nucleotide sequences and for inserting, changing or deleting such sequence into host sequences.

FROM MOLECULAR ERRORS TO INTERACTIONAL MOTIFS: RNA PARASITES AS OPEN SPACE INVADERS

In most origin of life scenarios, RNA parasites are considered major barriers for the origin of code and life (systems) which compels these proposals to close off the action of parasites into self contained code systems. But RNA parasites can also provide new and highly dynamic code that is added to the system. Naturally evolved RNA sequences can never be completely specified (or closed), since they must interact with their environment, replicate and undergo adaptation while retaining code that can always be further parasitized. Open systems can thus embrace the capacity of parasites to add novelty. This contrasts sharply with closed systems which must limit all such parasites.

This open feature renders the ability to absolutely specify membership (absolute immunity) as basically

indefinable. Any naturally evolved nucleotide sequence can never be fully secure from as yet undefined parasite agents. But a crucial inference out of this “insecurity” is that parasites provide the inherent capacity for novelty, *i.e.*, the precondition for greater complexity. Since parasites are competent in code, they are not “mistakes”. This means that the accepted core requirement of biological innovation, variation, is not met by the explanation model of random mutation (error) and its selection. Instead, RNA parasites provide continued infection and colonization which result in added identity of new RNA to existing RNA groups and thus alter self identity (immunity). These RNAs are thus acting as competent agents - not mutations/errors - which seek to impose new and competent code (identity) onto the system. Such RNAs, however, must interact to attain a stable colonization by inhibiting run-away parasite replication.

The core issue is thus to specify how RNA “agents” emerge from chemicals (ribozymes) to form the needed identity (such as for replicators) and also to form RNA groups that can support themselves and learn new membership.

For all extant life, these agents must have initially been RNA stem-loops. Single RNA stem-loop generation occurs by physical chemical properties solely as demonstrated by natural and randomized RNA experiments^[29-31]. If numbers of stem-loops are able to build complex consortia with greater competence, they would then represent the initial cooperative interactions needed to develop living systems that are not present in a strictly chemical world. The resulting system must function to maintain itself. In consortia, the emergence of identity (ability to differentiate self *vs* non-self) is a crucial initial step. Thus we seek to understand how single RNA stem-loop RNAs can become competent RNA consortia. And in so doing, we invoke the central action of RNA parasites and follow how “parasite-derived” RNA stem-loops interact in social collectives promoting innovation, infection, immunity and complex multiple (group) identity^[32].

THE CORRECTIVE AND COLLECTIVE POWER OF VIRUS VIA QUASISPECIES CONSORTIA (QS-C)

The term quasispecies originated from models that described related viral RNA populations resulting from error based variation of the master fittest type^[1-3]. It was not initially applied to consortia that showed cooperation. In the ensuing several decades, many laboratory observations were made that indicated more complex collective behaviours for viral quasispecies than were predicted by Eigen’s quasispecies equations. Two of the more active laboratories were those of John Holland and Esteban Domingo^[33]. The most recent compilation of these studies outlines many of the collective behaviours that have been made with quasispecies^[34]. The culmination study that most clearly reported that quasispecies have more complex collective behaviours seems to be the study from

the Andino group of poliovirus pathogenesis in a mouse model in which diversity and cooperation were key to viral fitness^[35,36]. Such studies led to the set of statements above on the cooperative nature of quasispecies. Thus quasispecies are collectives that have positive and negative interacting members that are bound together for a combined fitness that depends on diversity^[36-38]. It is thus ironic that it is from the viruses, assumed for decades to be the most selfish of all genetic entities, we observe the characteristics of cooperative, collective behaviour. And it was the “fittest type” assumptions of Manfred Eigen^[1] that generated quasispecies equations and theory which stimulated the development of this modern collective quasispecies view for over 40 years. But we are left with a conceptual contradiction. Modern quasispecies observations do not depend on the master (fittest type) and the consensus sequence. Consensus sequence may not predict the fitness of the diverse collective. In contrast to this diversity itself seems crucial.

QUASISPECIES-CONSORTIA PRODUCE HIGH RATES OF DIVERSITY: THIS IS NOT ERROR

With this clarification, it should become apparent that all RNA replicators (especially simple ones) must have high rates of diversity generation (not error). Novelty is then generated from new combinations of this diversity. Indeed, it sounds curious to use the term “error prone” for the high production rates of sequence novelties. An error is an inferior (“less better”) variant of an extant sequence. Such errors should only provide rare incremental improvement and be much less (if at all) able to generate *de novo* networks. With this error concept we also apply terms such as “damage”, “defect” and “incomplete” to variant information.

In contrast to this, cooperative RNA quasispecies produce and configure sequence novelties that are members of coherent populations and must generate an interacting diversity as prerequisite of innovation (variation), the driving forces of evolution. As an analogy, we might apply the limits of the “error” concept to innovative human endeavours. For example; poets produce novel poems by reconfiguring the commonly shared vocabulary. Must they also be error producers? Similarly, music composers and all artists in general produce novelty from common combinatorial rules and existing basic material tools. Do they also operate as error prone agents of fine arts? As we will see below, the qs-c concept requires interaction and diversity so it can even help us understand these processes of human innovation. But we must also apply the shared nature of “agents” that can create novelty. We will now consider how stem-loop RNAs can help us understand “agents” and their shared common use.

But in addition to stem-loop RNAs, all genetic entities that replicate via RNA will also be prone to qs-c (collective) behaviors. These behaviors will include cooperative and competitive interactions. RNA, however, is

not simply providing syntax for genetic information. It is more than code. It can also provide: (1) structure (stem-loop); (2) identity (stem-loops, 5', 3' ends); (3) functional (ribozyme, endonuclease, ligase) activity.

And it may be dynamic (*e.g.*, pseudoknots). Because of this extended capacity relative to DNA, RNA can be considered as a more active entity, with behaviours that make it able to function as an “agent” to affect its own activity and survival^[39,18]. At this point we adapt the framework from pure physics and chemistry to emergent biotic agent-based group building. In that light, DNA can be considered as a habitat for various RNA agents. It was from this perspective that we proposed that DNA should be considered as a habitat for these active RNA agents^[18,32]. But this discussion of simple RNA replicators suggests that the concept of qs-c should also apply to the ideas and experiments concerning the “RNA world” hypothesis. Yet curiously, very little “RNA world” research has addressed any issues regarding quasispecies^[40-42], let alone the more modern qs-c idea. As many are starting to think that life originated in a cooperating situation^[43], it is worth briefly considering if the qs-c concept will provide a different scenario for the origin of life.

THE RNA WORLD RECONSIDERED: INFECTIOUS STEM-LOOPS THAT OPERATE VIA QS-CONSORTIA

To evaluate the qs-c and infectious perspective on the RNA world hypothesis, we apply and explore the RNA-agent concept introduced above to the role of stem-loop ribozymes in the origin of life. The main objective is to incorporate the historically absent qs-c and parasitic perspective (with its inherent feature for group fitness) into the process that creates RNA societies. We will not explore early chemical evolution that might have led to the emergence of RNA molecules, but will instead assume RNA has come into existence and follow its features from this perspective.

One immediate consequence of this perspective is that we will be focused on collective features of RNA populations and will thus evaluate the chemical consequence of ribozyme qs societies, not individual replicators. This foundation immediately creates a situation in which “systems” of molecules with multiple behaviours will have the primary role in promoting the origin of life. It will also be important early on to consider how these systems maintain coherence (group identity, presented below), as this is an essential feature. Indeed, a basic and continuing theme will be that a core function of stem-loop RNAs is to provide molecular identity through all of evolution, including recent human evolution. This identity theme will persist throughout this chapter and will be frequently reintroduced.

The idea is then that individual members of stem-loop RNA societies were collectively able to invade (ligate into) each other to form a more stable and capable (ribozyme active) consortia with emergent, transformative

and unpredictable abilities. These collectives would lead to the origin of various ribosome- and RNA cellular societies (still linked to its stem-loop tRNA origin). Such a scenario also introduces the basic role of cooperation in the origin of life. It does not, however, eliminate competition, preclusion or extinction which are also inherent features of qs-c behaviors. Furthermore, the identity and transmissive role for stem-loop RNAs set the early (precellular) foundation for the origin of viruses whose emergence will further drive host evolution via colonization.

The cooperative and parasitic features of qs-c will also promote the early participation of peptides in the identity and evolution of the RNPs. The maintenance of these RNPs as a coherent collective will generally be mediated by addiction modules, which underlie group identity and immunity in all living systems. Addiction modules are counterbalanced (former competing) genetic parasites which share a persistent life style in host genomes. Addiction modules are clearly the result of stable consortial interactions^[4,44-46].

With this foundation, the emergence of genes, DNA, cells and individual fittest type selection can all be derived. But the emergence of DNA and cells and Darwinian evolution do not terminate the central role for transmissive RNA societies in the evolution of life. DNA becomes a habitat for these stem-loop “identity” RNAs and it is from this perspective that we can subsequently examine recent events in human evolution.

One issue should already be clear: This scenario posits that collective and cooperative behaviours were and remain essential for the emergence of living complexity. Qs-c then provides a conceptual foundation for the study of cooperating chemical biotic (in contrast to abiotic) networks in which mixtures of self-replicating RNA ribozymes can form highly cooperative and dynamic autocatalytic cycles^[29,31]. Let us now put this into the perspective of viroolution^[11].

TO SURVIVE RNA STEM-LOOP REPLICATORS MUST FORM POPULATIONS THAT DYNAMICALLY GENERATE IDENTITY

In the origin of the RNA world, short RNA oligomers formed by chemical processes needed to become longer RNAs able to perform template based catalysis. It has been proposed that the initial chemical formation of hairpin-like RNAs (stem-loops) could provide ribozyme activity following a ligation based modular evolution that would yield ribozyme auto catalysis^[47]. Indeed, below we present a series of studies that support this modular view.

But according to the parameters of qs-c evolution, for a consortium of RNA stem-loop replicators to survive, they must form a coherent population. They must share their identity and survival. The recognition of the stem-loop sequence itself by catalytic agents could provide such common identity. Alternatively, chemical markers or

initiators of catalysis could also mark the common population for priming or replication. Thus it is very interesting that the smallest ribozyme so far reported consists of just 5 nucleotides able to catalyze aminoacylation of the 3' end^[48].

The addition of an amino acid to an RNA molecule has many interesting chemical implications. A ribozyme has rather limited chemical potential compared to proteins. This is mostly due to proton disassociation constant of various amino acid moieties which are not close to pH neutrality. Thus amino acids are much more capable as chemical catalyst for this reason. Without the participation of amino acids, ribozymes must attain complex folds, often with some dynamic character (pseudoknots) to be effective catalyst allowing them to cleave and ligate RNA. Given this chemical advantage, we might expect that RNA evolution was greatly facilitated (but not coded) by peptides that contribute catalytically. In addition, such a modified RNA would likely also provide a chemical marker that could distinguish this RNA population. Indeed this molecular identity idea is developed below as a way to better understand the origin of tRNA and its role in initiating replication of so many RNA viruses, as well as how this chemical marker could promote the symbolic genetic code.

BASIC GROUP-BUILDING OF RIBOZYMES THROUGH SELF-LIGATION OF RNA STEM-LOOP MODULES

A good starting point for the accumulation of complexity seems to be hairpin ribozymes whose activity can be controlled by external effectors^[49]. Structural variation in these ribozymes allows progeny RNA to have different functions from their parental RNAs. The objective is to replicate RNA with RNA which hairpin ribozymes can perform as a sequence of ligation reactions that produce a longer ribozyme^[50]. Along these lines, two short hairpin RNAs can catalyze their own ligation to form larger RNA constructs^[50]. Thus we see interactions that promote more complex progeny. However, for a fully active ribozyme, complex RNA folding is needed. And such folding is cooperative^[51].

Folded ribozymes can also interact with other small molecules promoting their function as riboswitches^[52,53]. This includes amino acids which could promote either catalytic control or group identity marking. And the ribozyme folds can also be dynamic and context sensitive as seen in pseudoknots^[54].

But ribozymes can also be invasive, including self invasive^[55]. Thus stem-loop RNAs have many behaviors that would allow them to function as an identity group of agents involved in their own recognition and synthesis. Of particular interest is their ability to self-ligate^[50] as this could promote the emergence of RNA societies with self-identity. We can also think of tRNA as stem-loop RNA with various functions and histories. Indeed, it ap-

pears that tRNAs evolved from two separate hairpins^[56], in which each of the stem-loops interacts with a different ribosomal RNA subunit (presented below). This is a very interesting observation from an RNA society perspective. The invasive nature of intron ribozymes (endonuclease) also applies to tRNA from archaea, but here four distinct specificities are known^[57]. This very much resembles an identity system in which introns are marking central cellular (self) agents (tRNAs) for group identity but should destroy similar tRNAs (viral, *etc.*) lacking the intron marking. It is thus also interesting that tRNA with various linked amino acids themselves have been proposed to have originated before the translation system as genomic 3' tags needed for RNA ribozyme replication^[58-60]. This early function can also be explained as having served as a tag for group identity and could better explain the polyphyletic nature of the origin of tRNA^[61].

Interestingly viroids, the smallest virus-like agents, which infect plants, have striking similarities to Hepatitis Delta Virus^[62]. This virus, the smallest genome of any animal virus, uses circular genome, secondary structure folding and replication by a rolling circle mechanism that is catalyzed by host enzymes and cis-acting ribozymes. Mobile genetic elements are also typical self-splicing ribozymes. They excise their own RNA, from precursors thus support their own identity. The module like structure is evident. The viroids rely only on RNA and its structural motif, so it is really the sequence itself and its secondary structure which represents the entire infectious agent^[62].

THE ROLES OF RNA STEM-LOOP VARIATIONS IN QS-CONSORTIA BUILDING

This inherent capacity to form stem-loops with loops, bulges, junctions that are not immediately repaired or corrected opens the possibility to build abundance of varieties, which alter compositional patterns, identities, immunities and the whole row of progeny within a quasi-species-consortia (qs-c).

Let us consider a single stem-loop RNA that undergoes several rounds of replication. Potential identity effects of diversity may integrate the following: (1) altered self; (2) new interaction with other RNA; (3) complementing replication; (4) interfering with replication; (5) serving as primer for new replication pattern; (6) serving as target for ligase; (7) serving as target for cleavage; (8) serving as target for integration; (9) serving as initiation point of strand opening; and (10) serving as interaction point with peptide.

A variety of combinations of the above listed outcomes multiplies identity-generating and identity-shifting effects (Figure 1).

Each replication event - necessarily being low fidelity - produces its own peculiar version of diversified progeny, *e.g.*, with a new "bulge" in the stem. This bulge then becomes available to provide a whole array of possible

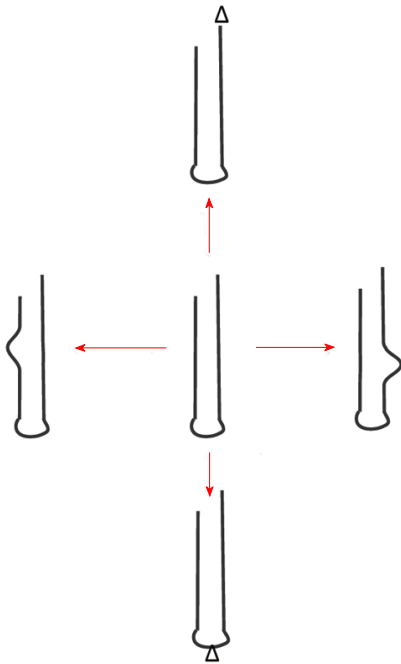


Figure 1 Self-generating RNA-stem-loop innovations: Generation of diversity (innovation) is shown. This is different from errors (accident, damage). Red arrow denotes diverse products from central template by low fidelity replication. Bulges and triangles denote changes.

outcomes (including contradicting ones).

It might: (1) interact with the original template to either compliment or inhibit it; (2) provide an interaction point for other RNA progeny (including itself); (3) provide a target site for cleavage or ligation; (4) act in combination with other progeny to provide more complex catalytic (ribozyme) function; and (5) alter or provide a binding site to other participants, such as peptides (RNPs).

In other words, even a singly new altered RNA now has a whole array of possible and multiple usages (positive and negative). Whole actual use will depend on the circumstances and history of the population it is in, *i.e.*, actual use depends on context.

If we add to this all the other diverse RNA progeny from these few rounds of replication - all in their own peculiar RNA region - all providing their own peculiar potential for use, we start to see the multiple potential for each individual RNA.

With such a scenario this combinations of possibilities very rapidly become too complex to follow the fate (fitness, usage) of any particular RNA. But this is the wrong, since it is an application of linear thinking. Instead, we need nonlinear thinking. Thus, if we think in sociological terms, then the RNA population (quasi species) can be considered as a “culture” that retains a common language which provides a level of group coherence (qs selection) on the basis of compatible cooperative organization. Each individual diverse RNA then becomes like a potentially new word for that language, *i.e.*, new agent in the ensemble of interacting agents.

The “culture” is then free to use a variant RNA (possibly even with multiple “meaning”), however it can also

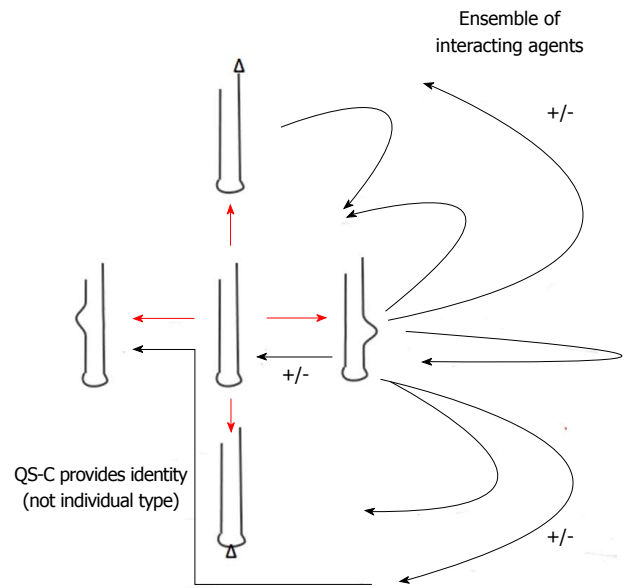


Figure 2 The crucial difference of quasispecies consortia with former quasispecies -concepts (fittest type – mutant spectra) is the basically consortial organisation of functional RNA ensembles. Shown above are the possible consortial interactions (black arrows) of just one the diversified RNA-stem-loops. These multiple activities (shown as +/-) preclude individual fitness definitions but require emergence and adaptation of group membership-identities. Defectives with similar subviral RNA-(stem-loop groups) remain relevant in both evolutionary and developmental processes. As a result of this basic evolutionary process of RNA stem-loop consortia building we can look at the emergence of *de novo* identities.

reject it. In this case, the term qs-c can have yet another meaning, c for culture (Figure 2).

These RNA uses will also vary considerable with the history of prior RNAs as well as any possible interactions with outer RNA quasi species. And the use can also vary (and be lost) with time as the culture adopts new meanings.

RNA-STEM-LOOPS AS MODULE-LIKE POOL

RNA stem-loops thus serve as multiple use and re-useable tools in RNA secondary sequence structure. Once being invented/generated such a single stem-loop may become part of another stem-loop consortium. This leads us not solely to the abundance of small and long non-coding RNAs that serve as regulatory tools in all known relevant cellular processes, but also to recombination processes that drive evolution^[63-68]. As a result we then can find a changed identity of the new consortium with altered features.

FROM RNAs TO RNAGENTS: FROM ABIOTIC CHEMICAL REACTIONS TO BIOTIC GROUP BEHAVIOUR

We have seen that from single stem-loop RNA populations of diverse stem-loops that acquire multifunctional “agent” behavior (RNAGents) can be generated. The

emergence of RNAs leads us to the emergence of life and life worlds. To understand the life-worlds of RNAs means to understand a fascinating sphere of what was formerly not assumed to be possible: The mere sequences of molecules do not behave like physical/chemical entities in an abiotic world but behave as competent agents on genetic code syntax that cooperate and organize, constitute and innovate sequence structures^[69,70]. Only groups of RNA stem-loops undergo such selection. Single RNA stem-loops behave like a random assembly of nucleotides without selective forces^[29,71,72]. Only if they are ligated to groups, *i.e.*, consortia, are they competent. This means they share a culture of interactional patterns and a history of defined time scales^[30,73]. Both culture and history are strongly influenced by ecosystem habitats that underlie selective forces and therefore are not reproducible features of abiotic matter.

The basic module of such RNAs are their complementary composition of base pairing stems and not-base pairing loops which results of an inherent property of RNA ensembles, the fold back to complementary polyRNAs, as demonstrated by the variety of ribozymes^[74]. Also the genomes of plus strand RNA viruses are able to form secondary and higher-order RNA structures^[75]. Interestingly, recent experiments demonstrated that single self-replicating RNAs are not as successful as cooperative ensembles and as a result cooperative RNA consortia outcompete selfish RNAs^[31]. Also, diversity in the backbone of RNA replicators seems essential to allow strand separation and replication^[76]. Thus group cooperation is key at the very origins of RNA societies (this is coherent with empirical knowledge about the emergence of natural codes/languages: they are essentially social group features not *solus ipse* results; see Witzany 1995)^[16].

Additionally, such RNA stem-loop consortia as found in long noncoding RNAs take multiple roles in RNA binding, DNA-binding, conformational switch and protein binding. Their modular nature results in functions such as control regimes, scaffolding to RNPs as well as guiding in target recognition within genomes^[53]. It is very clear now that non-coding RNAs build a complex layer that determines the regulation of all steps and substeps of gene control in complex organisms^[18].

SINGLE RNAs JOIN GROUPS AND NETWORKS: THE FEATURES

The origin of spontaneously cooperating networks of stem-loop RNA replicators can be understood from the qs-c perspective in which cooperation is the essential behavioral motif that outcompetes selfish behavioral patterns^[31]. Thus we see the emergence of networks at early stages in the evolution of life. In this review the term network will be used to include some distinct features, specifically network membership. Basically, for a network to be coherent and able to act collectively, it must limit membership to promote coordination. Otherwise it is simply a collection of uncoordinated agents and there will be no selection for maintaining the network

coherence or existence. If we are examining a network composed of stem-loop RNAs, it will be necessary for the individual RNAs to have some feature or behaviors that maintains membership such as RNA replication and recognition. This requires interaction. If only one type of RNA is supported (*e.g.*, high fidelity replication), there can be no complementation and complex function (*i.e.*, ribozyme) for the collective. A diversity of behavior and type will be essential. Recall however, that these RNAs act as agents in which various (multiple) behaviors will be possible even for the same sequence. This means there is diversity of interaction as well as diversity of type is inherent to the network. Thus overall interaction of an RNA agent with the collective must promote coherence and continued existence. What then are the features that promote continued existence (selection) for a network?

This does not require that only positive (*e.g.*, replication) interactions be supported. Negative interactions, including interference will also be needed. For example, highly efficient run-away replicons would overtake a qs-c and yield only one RNA type. Thus the qs would lose complementing functionality and would also consume all substrates if they were not regulated. This situation presents a problem in those habitats with limited substrates which likely is a very common state. Therefore, some level of self-regulation (negation) in the collective would promote the survival of the collective, especially if these RNAs could interact with the substrate in a regulatory, *e.g.*, riboswitch-like manner.

That efficient replicators become susceptible to parasitic replicators would provide an inherently spontaneous process of self-regulation. Yet the collective will still need to promote replication when it is favored. Accordingly, it becomes important for members of the collective to be subjected to both positive and negative self-regulation *via* RNA-RNA interactions. However, here too there must be some limits to self-regulation as the collective cannot tolerate overly active self-regulating members that will extinguish the collective. Thus we see that being a successful member of a collective requires many (and multiple) behaviors associated with it.

On top of that, as a qs-c replicates, these features will drift with time in a dynamic manner. In this context we can see that a random RNA stem-loop or a stem-loop RNA from a different qs collective would likely not be coherent with the other members of a particular qs.

A qs society is generally rather specific for its members. Group selection has already occurred in generating the qs-c. Indeed, as many experiments with RNA viruses infecting humans and animals have shown, a particular qs will exclude other qs of the same virus, such self exclusion can provide an origin for immune functions^[44,77]. Such society membership is also time dependent in that the serial passage of the same viral qs will usually result in subsequent qs that preclude prior individual members. This behavior has often been called a "Red Queen" behavior, but such a classical neo-Darwinian view does not incorporate or acknowledge the issue of group membership (qs-c coherence).

GROUP MEMBERSHIP IS THE PREREQUISITE FOR SELF/NON-SELF DIFFERENTIATION COMPETENCE

An important consequence out of this perspective is that the membership view allows us to understand the maintenance of minority types in the collective since these members can provide a needed but complementing catalytic control. Thus a qs society is a network that will naturally promote the emergence of membership, not the destruction of minority types. The important side effect is, that defective interfering agents can also contribute to membership control.

As previously proposed^[45,46], group membership can also be promoted by the combined action of toxic agents linked to antitoxic agents. A common version of a toxic agent is an endonuclease that will cleave sequences that are recognized as foreign. The antitoxin in this case prevents the action of the endonuclease (*e.g.*, *via* a bound protein or methylated base, dsRNA with another molecule, altered RNA fold, *etc.*). In this light, the endonuclease and ligation activities of stem-loop ribozymes are particularly interesting. A stem-loop ligase could provide a mechanism to recognize non-member stem-loop RNAs and destroy them by ligation.

Recall, however, that serial ligation can also be used to copy a stem-loop RNA. But such a situation has several very interesting implications. One of the problems with a society of stem-loop RNAs is that to attain their combined function, they need precise physical molecular placement relative to one another. This would normally require a high concentration dependence to counteract diffusion. By ligation, however, we could build a society of stem-loop RNAs that have covalently placed the various stem-loops in the correct functional (or dynamic/regulatory) context and have lost their concentration dependence. It seems likely that such a process would involve invasive self-colonizing stem-loop RNAs that results in one molecular entity with a common identity function. This would generate one entity that evolved from the ligation of a mixed set of stem-loop agents that now have a highly enhanced (collective) functional capacity. This collective would also have a highly enhanced capacity for persistence as it need not continually replicate individual stem-loop RNA agents to maintain its membership. The collective, however, would still need to oppose non-member or other parasite participation. Additionally, a collective might attain a conditional (regulated) replication capacity if it incorporates stem-loop RNA riboswitches. It is by such a process that we can now consider the origin of the ribosome.

BEFORE AND AFTER RIBOSOME EMERGENCE

A big problem with thinking that viruses are essential agents for the emergence of life, however, is the ribosome

ome^[4,9,78-81]. The ribosome really defines the cell from virus and seemed to preclude virus from early evolution of the cell^[80]. It is now clear also that the ribosome acts as ribozyme^[82]. Yet the ribosome itself is an ensemble of two rather complex stem-loop societies of riboagents^[83], an ensemble that became “set” with the invention of DNA^[84]. Therefore it makes sense now to re-evaluate the RNA virus first hypothesis^[4] from the before ribosomal world (BR) and the after ribosomal world (AR). Of course, common themes of consortia, symbolic code, quasispecies, addiction modules, group identity, membership agents would provide the themes that could link the great BR-AR divide. Briones *et al*^[47] developed an interesting four step model on this in scenarios of evolution on both mineral surfaces and inside vesicles, such as: (1) abiotic polymerization of RNA oligomers, (2) folding of the RNA oligomers and ubiquity of hairpin structures, (3) ligation based modular evolution of RNA and finally (4) template-dependent RNA polymerase. This unites the history and culture-dependent derived stem-loops integrated in the two ribosomal subunits^[85].

MEMBERSHIP IS CRUCIAL FOR LIVING NETWORKS (SYSTEMS) TO EMERGE

In examining the literature relevant to qs, the RNA world and RNA network formation, we can indeed find some experimental evidence that supports qs and the spontaneous emergence of RNA networks. But almost completely lacking from such experiments is any evaluation of the membership issue. For example, quasispecies-like behavior has been observed with *in vitro* RNA replicator studies^[86]. Non-enzymatic template (peptide) directed autocatalytic systems can show network behavior^[87] and communities RNA ribozyme replicator sets can also show lateral evolution^[88]. Also rule-based computing simulations have been applied to similar systems in an effort to understand the emergence of parasites and antiparasites^[6]. Along these lines, the hypercycle kinetic model was proposed to be a system of cross catalyzing RNA replicators which depend on cooperation for growth^[3], but this is not a collective autocatalytic system as proposed above^[89]. Hypercycles as proposed are not able to tolerate parasites, let alone depend on them for development. Yet the biggest problem of all such studies is that there is no assumption regarding the basic importance of network or group membership.

Without this network membership concept and its attending dynamic strategies, authentic collective action does not emerge, systems do not develop. The dynamic nature of network membership and collective action pose many unsolved problems for existing qs theory. For example, how is the multi-potential of an individual RNA to be evaluated within the qs-consortia if we cannot specify all the other interactions and how they change with time? We cannot apply our current ideas of fitness to this individual RNA as the historical and population

context is key. Network membership needs to be prominently considered if we are to understand the origin of the ribosome and the genetic code.

Replicator identity marking *via* 3' aminoacylated of stem-loop RNAs appears most able to explain the origin of a tRNA mediated genetic code. For in contrast to Darwinian evolution, network members will generally have multiple ancestral histories. These members will mostly originate from separate parasitic lineages that were able to penetrate defenses and join the network sometimes in mixtures. They don't need to descend from one individual or even be from the same type of agent (virus, transposon, intron, intene, *etc.*).

From this perspective we can understand why the two halves of tRNA have distinct evolutionary histories, yet tRNA is a core agent for evolution of life. Thus neither the amino acid based (peptide) ancestors nor the RNA based ancestors need a common origin to participate in a symbiogenic network. Our qs-c concept supports such a network process and - additionally - network membership provides the basis for examining noncoding RNA based regulation needed for multicellular complexity^[90].

VIRAL CORE COMPETENCIES: INNOVATION, INTEGRATION, REGULATION, EXAPTATION

Because infection derived domestications, such as the whole variety of retroposon derived non-coding RNA, are now known to be transcribed and to control the regulation of genes^[91], these parasitic agents shape genome architecture and function-arrangements. We now realize life on planet Earth has always resided within a virosphere and that the evolution of species depended on the virosphere features of innovation and transfer^[4,39]. Humans share similar gene number as *C.elegans*. And humans and mice share 98.5% coding DNA. Regulatory complexity thus seems to characterize the evolution of more complex eukaryotes and stems from parasitic elements. The remaining 1.5% protein coding information is differently regulated *via* these species specific non-coding elements in increasingly complex ways. These elements act together through a variety of combinations, situations and mechanisms to re-shape genome/gene functions through epigenetic imprinting and re-printing^[23]. These regulatory RNAs mostly retain the stem-loop structure. Also interesting are reports that ingestion of small RNA gene regulators seems to be usual route of RNA transfer^[92].

In this perspective we have applied qs-consortia concept to explain complex regulatory network formation in the origin of life and the cellular protein world of higher organisms. The real species that determine all these evolutionary patterns are viruses and virus-like (infectious) stem-loop RNA qs-consortia^[32,93,94].

Viral core competencies necessitates that they be competent in host genetic and epigenetic code. These competencies provide innovation, integration, regulation,

exaptation of qs consortia to form networks in the host.

A very intriguing example of this is found in placentation of mammals^[10,44]. From day 1-6 post fertilization, all vertebrate embryos are similar and divide to the morula stage. But on day 6 this morula hatches and becomes a trophoblast. From here on, embryo development differs significantly. The outer cell layer is the first committed tissue of the embryo and will become the placenta. Viruses have a real affinity (tissue specificity) for this layer. This is exactly where very high endogenous retrovirus (ERV) activity is found^[95-97]. This tissue has been repurposed to invade the uterine wall, suppress mother's (host) immune response, promote blood feeding (exchange), and alter mother's behaviour and physiology and brain. About 1500 placental genes are thought to have been modified by altered (ERV mediated) network re-regulation^[98]. In our own view, it was likely that a collection of ERVs were involved. Other viral agents associated with reproductive biology that we don't yet understand, were also likely to have been involved.

Consider the example of HIV-1 as a virus that requires qs diversity and dynamics to solve very complex problems and dynamic situations. If we think of HIV-1 as a fittest type agent with only about 10000 bp of clonal RNA, such a "pure" virus could not defeat the complex human adaptive and innate immune systems. An HIV-1 limited to the "master copy" fittest type (consensus sequence from a successful human infection) would be unable to generate a qs-c, and would rapidly be eliminated by our immune system and pose no problem for vaccine development. Here, qs-c is providing fitness. It is this power of the consortia to defeat complex systems that when applied to stem-loop RNAs can also provide the power for the generation of *de novo* diversity and cooperation needed to originate life.

CONCLUSION

In our expanding perspective from physics and chemistry to sociology, from elements/chemicals to emergence of agents, *i.e.*, from RNAs to RNAs, we found several indices that lead us to a new concept of RNA quasispecies. In contrast to former opinions and concepts that proposed single fittest type and its mutant spectra as the mechanisms of variation that drives evolution, qs-consortia depend on diversity, multiuse (counter) active agents and consortia membership. Thus, when we think of a commonly shared genetic code that is used and represented by consortia of RNA stem-loops, it is providing not only information storage but active group membership-identity. Therefore we term this membership-identity qs-c. In this perspective qs-RNA and virus evolution are inherently cooperative and modular. The essential players are not fittest types, but consortia of RNAs that need diversity.

The basic rule here is that of innovation and group selection. The modular character of consortia building with its inherent self-ligation capability of basic stem-

loop tools is based additionally on the participation and integration of defectives/mutants (errors/junk) as energy optimizing resource for module-like re-usage. In such dynamic RNA stem-loop populations several basic behavioral motifs are combined that are absent in pre-biotic chemistry; such as complementation, cooperation, competition (selection), preclusion and lethal defections. This broad range of behaviors require a diversity that also integrates the remaining “memory” (information-storing) parts of essential minorities. Under changing contexts minorities may have previously been or may yet become majorities. Thus memory of past experience is inherent in and used by qs-c.

In no other natural language are the agents that communicate (coordinate and organize) *via* repertoire of natural signs (language) also identical with the signs (words) themselves. This is precisely what we have proposed with stem-loop RNAs. This proposition defines a new phenomenon: at the beginning of life agents and “words” (information) are identical. What has been divided since invention of DNA and LUCAs (signs from agents), was formerly unified. The qs-c sociology thus describes this unified status and its interactions in their current DNA/protein habitats. Indeed it is difficult to formulate sentences about a status that never have been formulated before: Agents that represent sequences of signs are themselves subject of sequence generation (as described above). In the current RNA world (now residing in their DNA habitats) they are still alive. The contrast to the early RNA world, in that their available habitats (DNA/protein) have expanded indefinitely.

Now we can imagine the move from RNA physics/chemistry to RNA sociology. Although they consist of atoms that bind together by laws of physics and chemistry they don't behave like abiotic ensembles but as semiotic subjects absent in abiotic world. “Words” are (stem-loop) agents and “sentences” are consortia of such agents. For example, some “sentences” result in tRNAs and the ribosomal subunits which assemble to ribosomes. Some result in RNA viruses, some in defectives RNAs that serve as effective non-coding RNAs in a regulation processes of host genomes. The high mutation rate is now recognized as freedom (from mechanistic determinism) to generate new sequence (consortia) space. “Error” now becomes innovation-competence for new generations of both RNA agents and biological information with unpredictable competence and membership.

ACKNOWLEDGMENTS

The authors are grateful to Professor Vera Kolb and to the reviewers for their constructive suggestions.

REFERENCES

- 1 **Eigen M.** Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* 1971; **58**: 465-523 [PMID: 4942363]
- 2 **Eigen M, McCaskill J, Schuster P.** The Molecular Quasi-

- species. *Adv Chem Phys* 1988; **75**: 149-263 [DOI: 10.1021/j100335a010]
- 3 **Eigen M.** Viral quasispecies. *Sci Am* 1993; **269**: 42-49 [PMID: 8337597]
- 4 **Villarreal LP.** Viruses and the Evolution of Life. Washington: ASM Press, 2005
- 5 **Forterre P.** The origin of viruses and their possible roles in major evolutionary transitions. *Virus Res* 2006; **117**: 5-16 [PMID: 16476498]
- 6 **Jalasvuori M, Jalasvuori MP, Bamford JK.** Dynamics of a laterally evolving community of ribozyme-like agents as studied with a rule-based computing system. *Orig Life Evol Biosph* 2010; **40**: 319-334 [PMID: 20012776 DOI: 10.1007/s11084-009-9187-0]
- 7 **Witzany G (ed).** Viruses: Essential Agents of Life. Dordrecht: Springer, 2012
- 8 **Koonin EV, Senkevich TG, Dolja VV.** The ancient Virus World and evolution of cells. *Biol Direct* 2006; **1**: 29 [PMID: 16984643]
- 9 **Koonin EV.** On the origin of cells and viruses: primordial virus world scenario. *Ann N Y Acad Sci* 2009; **1178**: 47-64 [PMID: 19845627 DOI: 10.1111/j.1749-6632.2009.04992.x]
- 10 **Villarreal LP.** Origin of Group Identity: Viruses, addiction and cooperation. New York: Springer, 2009
- 11 **Ryan F.** Virolution. London: Harper Collins, 2009
- 12 **Roossinck MJ.** Lifestyles of plant viruses. *Philos Trans R Soc Lond B Biol Sci* 2010; **365**: 1899-1905 [PMID: 20478885 DOI: 10.1098/rstb.2010.0057]
- 13 **Roossinck MJ.** The good viruses: viral mutualistic symbioses. *Nat Rev Microbiol* 2011; **9**: 99-108 [PMID: 21200397 DOI: 10.1038/nrmicro2491]
- 14 **Roossinck MJ.** Persistent Plant Viruses: Molecular Hitchhikers or Epigenetic Elements? In: Witzany G (ed). Viruses: Essential Agents of Life. Dordrecht: Springer, 2012: 177-186
- 15 **Koonin EV.** The Logic of Chance. The Nature and Origin of Biological Information. New Jersey: FT Press, 2011
- 16 **Witzany G.** From the “logic of the molecular syntax” to molecular pragmatism. Explanatory deficits in Manfred Eigen's concept of language and communication. *Evol Cog* 1995; **1**: 148-168
- 17 **Shapiro JA.** Evolution: A view from the 21st century. Washington: FT Press, 2011
- 18 **Witzany G.** The agents of natural genome editing. *J Mol Cell Biol* 2011; **3**: 181-189 [PMID: 21459884 DOI: 10.1093/jmcb/mjr005]
- 19 **Witzany G, Baluška F.** Life's code script does not code itself. The machine metaphor for living organisms is outdated. *EMBO Rep* 2012; **13**: 1054-1056 [PMID: 23146891 DOI: 10.1038/embor.2012.166]
- 20 **Wittgenstein L.** Philosophical Investigations. Oxford: Blackwell Publishing, 2001
- 21 **Slotkin RK, Martienssen R.** Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* 2007; **8**: 272-285 [PMID: 17363976]
- 22 **Jirtle RL.** Epigenome: the program for human health and disease. *Epigenomics* 2009; **1**: 13-16 [PMID: 22122631 DOI: 10.2217/epi.09.16]
- 23 **Barlow DP.** Genomic imprinting: a mammalian epigenetic discovery model. *Annu Rev Genet* 2011; **45**: 379-403 [PMID: 21942369 DOI: 10.1146/annurev-genet-110410-132459]
- 24 **Briones C, de Vicente A, Molina-París C, Domingo E.** Minority memory genomes can influence the evolution of HIV-1 quasispecies in vivo. *Gene* 2006; **384**: 129-138 [PMID: 17059869]
- 25 **Briones C, Domingo E.** Minority report: hidden memory genomes in HIV-1 quasispecies and possible clinical implications. *AIDS Rev* 2008; **10**: 93-109 [PMID: 18615120]
- 26 **Bordería AV, Lorenzo-Redondo R, Pernas M, Casado C, Alvaro T, Domingo E, Lopez-Galindez C.** Initial fitness recovery of HIV-1 is associated with quasispecies heterogene-

- ity and can occur without modifications in the consensus sequence. *PLoS One* 2010; **5**: e10319 [PMID: 20436678 DOI: 10.1371/journal.pone.0010319]
- 27 **Grande-Pérez A**, Lázaro E, Lowenstein P, Domingo E, Manrubia SC. Suppression of viral infectivity through lethal defection. *Proc Natl Acad Sci USA* 2005; **102**: 4448-4452 [PMID: 15767582]
- 28 **Domingo E**, Gomez J. Quasispecies and its impact on viral hepatitis. *Virus Res* 2007; **127**: 131-150 [PMID: 17349710]
- 29 **Smit S**, Yarus M, Knight R. Natural selection is not required to explain universal compositional patterns in rRNA secondary structure categories. *RNA* 2006; **12**: 1-14 [PMID: 16373489]
- 30 **Gwiazda S**, Salomon K, Appel B, Müller S. RNA self-ligation: from oligonucleotides to full length ribozymes. *Biochimie* 2012; **94**: 1457-1463 [PMID: 22465106 DOI: 10.1016/j.biochi.2012.03.015]
- 31 **Vaidya N**, Manapat ML, Chen IA, Xulvi-Brunet R, Hayden EJ, Lehman N. Spontaneous network formation among cooperative RNA replicators. *Nature* 2012; **491**: 72-77 [PMID: 23075853 DOI: 10.1038/nature11549]
- 32 **Villarreal LP**, Witzany G. The DNA Habitat and its RNA Inhabitants. At the dawn of RNA sociology. *Genom Ins* 2013; **6**: 1-12 [DOI: 10.4137/GEI.S11490]
- 33 **Domingo E**, Parrish CR, Holland JJ. Origin and Evolution of Viruses, 2nd Edition. San Diego: Academic Press, 2008
- 34 **Domingo E**, Sheldon J, Perales C. Viral quasispecies evolution. *Microbiol Mol Biol Rev* 2012; **76**: 159-216 [PMID: 22688811 DOI: 10.1128/MMBR.05023-11]
- 35 **Vignuzzi M**, Stone JK, Arnold JJ, Cameron CE, Andino R. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* 2006; **439**: 344-348 [PMID: 16327776]
- 36 **Lauring AS**, Andino R. Quasispecies theory and the behavior of RNA viruses. *PLoS Pathog* 2010; **6**: e1001005 [PMID: 20661479 DOI: 10.1371/journal.ppat.1001005]
- 37 **Ojosnegros S**, Perales C, Mas A, Domingo E. Quasispecies as a matter of fact: viruses and beyond. *Virus Res* 2011; **162**: 203-215 [PMID: 21945638 DOI: 10.1016/j.virusres.2011.09.018]
- 38 **Arbiza J**, Mirazo S, Fort H. Viral quasispecies profiles as the result of the interplay of competition and cooperation. *BMC Evol Biol* 2010; **10**: 137 [PMID: 20459709 DOI: 10.1186/1471-2148-10-137]
- 39 **Witzany G.** (ed) Natural Genetic Engineering and Natural Genome Editing. New York: Annals of the New York Academy of Sciences, 2009
- 40 **Gesteland RF**, Cech T, Atkins JF. The RNA World: the Nature of Modern RNA Suggests a Prebiotic RNA. Vol. 2nd. Cold Spring Harbor Monograph Series, 37. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1999
- 41 **Atkins JF**, Gesteland RF, Cech Tr, (editors). RNA Worlds, New York: Cold Spring Harbor Laboratory Press, 2010
- 42 **Altman S.** Rna008 - BioTheory. 2013. Available from: URL: <https://sites.google.com/site/theoreticalbiology/main-results/rna-world/rna008>
- 43 **Holmes B.** First Life May Have Survived by Cooperating. *New Scientist* 2012; **216**: 10 [DOI: 10.1016/S0262-4079(12)62673-6]
- 44 **Villarreal LP.** The source of self: genetic parasites and the origin of adaptive immunity. *Ann N Y Acad Sci* 2009; **1178**: 194-232 [PMID: 19845639 DOI: 10.1111/j.1749-6632.2009.05020.x]
- 45 **Villarreal L.** Viruses and host evolution: virus-mediated self identity. *Adv Exp Med Biol* 2012; **738**: 185-217 [PMID: 22399381 DOI: 10.1007/978-1-4614-1680-7_12]
- 46 **Villarreal LP.** The Addiction Module as a Social Force. In: Viruses: Essential Agents of Life. Witzany G (ed). Dordrecht: Springer 2012: 107-145
- 47 **Briones C**, Stich M, Manrubia SC. The dawn of the RNA World: toward functional complexity through ligation of random RNA oligomers. *RNA* 2009; **15**: 743-749 [PMID: 19318464 DOI: 10.1261/rna.1488609]
- 48 **Turk RM**, Chumachenko NV, Yarus M. Multiple translational products from a five-nucleotide ribozyme. *Proc Natl Acad Sci USA* 2010; **107**: 4585-4589 [PMID: 20176971 DOI: 10.1073/pnas.0912895107]
- 49 **Müller S**, Appel B, Krellenberg T, Petkovic S. The many faces of the hairpin ribozyme: structural and functional variants of a small catalytic RNA. *IUBMB Life* 2012; **64**: 36-47 [PMID: 22131309 DOI: 10.1002/iub.575]
- 50 **Cheng LK**, Unrau PJ. Closing the circle: replicating RNA with RNA. *Cold Spring Harb Perspect Biol* 2010; **2**: a002204 [PMID: 20554706 DOI: 10.1101/cshperspect.a002204]
- 51 **Behrouzi R**, Roh JH, Kilburn D, Briber RM, Woodson SA. Cooperative tertiary interaction network guides RNA folding. *Cell* 2012; **149**: 348-357 [PMID: 22500801 DOI: 10.1016/j.cell.2012.01.057]
- 52 **Ferré-D'Amaré AR.** Use of a coenzyme by the glmS ribozyme-ribo-switch suggests primordial expansion of RNA chemistry by small molecules. *Philos Trans R Soc Lond B Biol Sci* 2011; **366**: 2942-2948 [PMID: 21930586 DOI: 10.1098/rstb.2011.0131]
- 53 **Mercer TR**, Mattick JS. Structure and function of long non-coding RNAs in epigenetic regulation. *Nat Struct Mol Biol* 2013; **20**: 300-307 [PMID: 23463315 DOI: 10.1038/nsmb.2480]
- 54 **Perreault J**, Weinberg Z, Roth A, Popescu O, Chartrand P, Ferbeyre G, Breaker RR. Identification of hammerhead ribozymes in all domains of life reveals novel structural variations. *PLoS Comput Biol* 2011; **7**: e1002031 [PMID: 21573207 DOI: 10.1371/journal.pcbi.1002031]
- 55 **Kumar RM**, Joyce GF. A modular, bifunctional RNA that integrates itself into a target RNA. *Proc Natl Acad Sci USA* 2003; **100**: 9738-9743 [PMID: 12913125]
- 56 **Dick TP**, Schamel WA. Molecular evolution of transfer RNA from two precursor hairpins: implications for the origin of protein synthesis. *J Mol Evol* 1995; **41**: 1-9 [PMID: 7608982]
- 57 **Fujishima K**, Sugahara J, Miller CS, Baker BJ, Di Giulio M, Takesue K, Sato A, Tomita M, Banfield JF, Kanai A. A novel three-unit tRNA splicing endonuclease found in ultrasmall Archaea possesses broad substrate specificity. *Nucleic Acids Res* 2011; **39**: 9695-9704 [PMID: 21880595 DOI: 10.1093/nar/gkr692]
- 58 **Sun FJ**, Fleurdépine S, Bousquet-Antonelli C, Caetano-Anollés G, Deragon JM. Common evolutionary trends for SINE RNA structures. *Trends Genet* 2007; **23**: 26-33 [PMID: 17126948]
- 59 **Sun FJ**, Caetano-Anollés G. Evolutionary patterns in the sequence and structure of transfer RNA: early origins of archaea and viruses. *PLoS Comput Biol* 2008; **4**: e1000018 [PMID: 18369418 DOI: 10.1371/journal.pcbi.1000018]
- 60 **Rodin AS**, Szathmáry E, Rodin SN. On origin of genetic code and tRNA before translation. *Biol Direct* 2011; **6**: 14 [PMID: 21342520 DOI: 10.1186/1745-6150-6-14]
- 61 **Di Giulio M.** A polyphyletic model for the origin of tRNAs has more support than a monophyletic model. *J Theor Biol* 2013; **318**: 124-128 [PMID: 23174278 DOI: 10.1016/j.jtbi.2012.11.012]
- 62 **Flores R**, Ruiz-Ruiz S, Serra P. Viroids and hepatitis delta virus. *Semin Liver Dis* 2012; **32**: 201-210 [PMID: 22932968 DOI: 10.1055/s-0032-1323624]
- 63 **Batzer MA**, Deininger PL. Alu repeats and human genomic diversity. *Nat Rev Genet* 2002; **3**: 370-379 [PMID: 11988762]
- 64 **Eickbush TH**, Jamburuthugoda VK. The diversity of retrotransposons and the properties of their reverse transcriptases. *Virus Res* 2008; **134**: 221-234 [PMID: 18261821 DOI: 10.1016/j.virusres.2007.12.010]
- 65 **Feschotte C.** Transposable elements and the evolution of regulatory networks. *Nat Rev Genet* 2008; **9**: 397-405 [PMID: 18368054 DOI: 10.1038/nrg2337]

- 66 **Klenerman P**, Hengartner H, Zinkernagel RM. A non-retroviral RNA virus persists in DNA form. *Nature* 1997; **390**: 298-301 [PMID: 9384383]
- 67 **Mahajan VS**, Drake A, Chen J. Virus-specific host miRNAs: antiviral defenses or promoters of persistent infection? *Trends Immunol* 2009; **30**: 1-7 [PMID: 19059006 DOI: 10.1016/j.it.2008.08.009]
- 68 **Schmitz J**, Brosius J. Exonization of transposed elements: A challenge and opportunity for evolution. *Biochimie* 2011; **93**: 1928-1934 [PMID: 21787833 DOI: 10.1016/j.biochi.2011.07.014]
- 69 **Qureshi IA**, Mehler MF. Emerging roles of non-coding RNAs in brain evolution, development, plasticity and disease. *Nat Rev Neurosci* 2012; **13**: 528-541 [PMID: 22814587 DOI: 10.1038/nrn3234]
- 70 **Witzany G**. From Molecular Entities to Competent Agents: Viral Infection-derived Consortia act as Natural Genetic Engineers. In: Witzany, G (ed). *Viruses: Essential Agents of Life*. Dordrecht, Springer. 2012 b: 407-419
- 71 **Gevertz J**, Gan HH, Schlick T. In vitro RNA random pools are not structurally diverse: a computational analysis. *RNA* 2005; **11**: 853-863 [PMID: 15923372]
- 72 **Przybilski R**, Hammann C. The tolerance to exchanges of the Watson Crick base pair in the hammerhead ribozyme core is determined by surrounding elements. *RNA* 2007; **13**: 1625-1630 [PMID: 17666711]
- 73 **Venner S**, Feschotte C, Biéumont C. Dynamics of transposable elements: towards a community ecology of the genome. *Trends Genet* 2009; **25**: 317-323 [PMID: 19540613 DOI: 10.1016/j.tig.2009.05.003]
- 74 **Moore PB**. The RNA Folding Problem. In: *The RNA World*, Second Edition. New York: Cold Spring Harbor Laboratory Press 1999: 381-401
- 75 **Liu Y**, Wimmer E, Paul AV. Cis-acting RNA elements in human and animal plus-strand RNA viruses. *Biochim Biophys Acta* 2009; **1789**: 495-517 [PMID: 19781674 DOI: 10.1016/j.bbagr.2009.09.007]
- 76 **Hernández AR**, Piccirilli JA. Chemical origins of life: Prebiotic RNA unstuck. *Nat Chem* 2013; **5**: 360-362 [PMID: 23609081 DOI: 10.1038/nchem.1636]
- 77 **Villarreal LP**. Viral ancestors of antiviral systems. *Viruses* 2011; **3**: 1933-1958 [PMID: 22069523 DOI: 10.3390/v3101933]
- 78 **Brüssow H**. The quest for food: a natural history of eating. Berlin/Heidelberg/New York: Springer, 2007
- 79 **Forterre P**. The two ages of the RNA world, and the transition to the DNA world: a story of viruses and cells. *Biochimie* 2005; **87**: 793-803 [PMID: 16164990]
- 80 **Forterre P**, Prangishvili D. 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* 2009; **1178**: 65-77 [PMID: 19845628 DOI: 10.1111/j.1749-6632.2009.04993.x]
- 81 **Jalasvuori M**. Revolutionary Struggle for Existence: Introduction to Four Intriguing Puzzles in Virus Research. In: Witzany G (ed). *Viruses: Essential Agents of Life*. Dordrecht: Springer 2012: 1-19
- 82 **Cech TR**. Structural biology. The ribosome is a ribozyme. *Science* 2000; **289**: 878-879 [PMID: 10960319]
- 83 **Eickbush TH**, Eickbush DG. Finely orchestrated movements: evolution of the ribosomal RNA genes. *Genetics* 2007; **175**: 477-485 [PMID: 17322354]
- 84 **Harish A**, Caetano-Anollés G. Ribosomal history reveals origins of modern protein synthesis. *PLoS One* 2012; **7**: e32776 [PMID: 22427882 DOI: 10.1371/journal.pone.0032776]
- 85 **Bokov K**, Steinberg SV. A hierarchical model for evolution of 23S ribosomal RNA. *Nature* 2009; **457**: 977-980 [PMID: 19225518 DOI: 10.1038/nature07749]
- 86 **Díaz Arenas C**, Lehman N. Quasispecies-like behavior observed in catalytic RNA populations evolving in a test tube. *BMC Evol Biol* 2010; **10**: 80 [PMID: 20331885 DOI: 10.1186/1471-2148-10-80]
- 87 **Dadon Z**, Wagner N, Cohen-Luria R, Ashkenasy G. 2012. Reaction Networks. In: Gale P, Steed J, editors. *Supramolecular Chemistry: From Molecules to Nanomaterials*. John Wiley and Sons, Inc, 2012 [DOI: 10.1002/9780470661345.smc162]
- 88 **Hordijk W**, Steel M. A Formal Model of Autocatalytic Sets Emerging in an RNA Replicator System. *J Syst Chem* 2013, **4**: 3 [DOI: 10.1186/1759-2208-4-3]
- 89 **Szathmáry E**. On the Propagation of a Conceptual Error Concerning Hypercycles and Cooperation. *J Syst Chem* 2013; **4**: 1-4 [DOI: 10.1186/1759-2208-4-1]
- 90 **Lozada-Chávez I**, Stadler PF, Prohaska SJ. "Hypothesis for the modern RNA world": a pervasive non-coding RNA-based genetic regulation is a prerequisite for the emergence of multicellular complexity. *Orig Life Evol Biosph* 2011; **41**: 587-607 [PMID: 22322874 DOI: 10.1007/s11084-011-9262-1]
- 91 **Mattick JS**, Taft RJ, Faulkner GJ. A global view of genomic information—moving beyond the gene and the master regulator. *Trends Genet* 2010; **26**: 21-28 [PMID: 19944475 DOI: 10.1016/j.tig.2009.11.002]
- 92 **Zhang L**, Hou D, Chen X, Li D, Zhu L, Zhang Y, Li J, Bian Z, Liang X, Cai X, Yin Y, Wang C, Zhang T, Zhu D, Zhang D, Xu J, Chen Q, Ba Y, Liu J, Wang Q, Chen J, Wang J, Wang M, Zhang Q, Zhang J, Zen K, Zhang CY. Exogenous plant miR168a specifically targets mammalian LDLRAP1: evidence of cross-kingdom regulation by microRNA. *Cell Res* 2012; **22**: 107-126 [PMID: 21931358 DOI: 10.1038/cr.2011.158]
- 93 **Brosius J**. The contribution of RNAs and retroposition to evolutionary novelties. *Genetica* 2003; **118**: 99-116 [PMID: 12868601]
- 94 **Villarreal LP**, Witzany G. Viruses are essential agents within the roots and stem of the tree of life. *J Theor Biol* 2010; **262**: 698-710 [PMID: 19833132 DOI: 10.1016/j.jtbi.2009.10.014]
- 95 **Varela M**, Spencer TE, Palmarini M, Arnaud F. Friendly viruses: the special relationship between endogenous retroviruses and their host. *Ann N Y Acad Sci* 2009; **1178**: 157-172 [PMID: 19845636 DOI: 10.1111/j.1749-6632.2009.05002.x]
- 96 **Conley AB**, Jordan IK. Endogenous Retroviruses and the Epigenome. In: Witzany G (ed). *Viruses: Essential Agents of Life*. Dordrecht: Springer, 2012: 309-323
- 97 **Perot P**, Bolze PA, Mallet F. From Viruses to Genes: Syncytins. In: Witzany G (ed). *Viruses: Essential Agents of Life*. Dordrecht: Springer, 2012: 325-361
- 98 **Chuong EB**, Rumi MA, Soares MJ, Baker JC. Endogenous retroviruses function as species-specific enhancer elements in the placenta. *Nat Genet* 2013; **45**: 325-329 [PMID: 23396136 DOI: 10.1038/ng.2553]

P-Reviewer Takahashi H S-Editor Song XX L-Editor A
E-Editor Wu HL