

Evolvability Is an Evolved Ability: The Coding Concept as the Arch-Unit of Natural Selection

Srdja Janković¹ · Milan M. Ćirković^{2,3}

Received: 7 June 2015 / Accepted: 11 August 2015 /

Published online: 29 September 2015

© Springer Science+Business Media Dordrecht 2015

Abstract Physical processes that characterize living matter are qualitatively distinct in that they involve encoding and transfer of specific types of information. Such information plays an active part in the control of events that are ultimately linked to the capacity of the system to persist and multiply. This algorithmicity of life is a key prerequisite for its Darwinian evolution, driven by natural selection acting upon stochastically arising variations of the encoded information. The concept of evolvability attempts to define the total capacity of a system to evolve new encoded traits under appropriate conditions, i.e., the accessible section of total morphological space. Since this is dependent on previously evolved regulatory networks that govern information flow in the system, evolvability itself may be regarded as an evolved ability. The way information is physically written, read and modified in living cells (the “coding concept”) has not changed substantially during the whole history of the Earth’s biosphere. This biosphere, be it alone or one of many, is, accordingly, itself a product of natural selection, since the overall evolvability conferred by its coding concept (nucleic acids as information carriers with the “rulebook of meanings” provided by codons, as well as all the subsystems that regulate various conditional information-reading modes) certainly played a key role in enabling this biosphere to survive up to the present, through alterations of planetary conditions, including at least five catastrophic events linked to major mass extinctions. We submit that, whatever the actual prebiotic physical and chemical processes may have been on our home planet, or may, in principle, occur at some time and place in the Universe, a particular coding concept, with its respective potential to give rise to a biosphere, or class of

✉ Milan M. Ćirković
mcirkovic@aob.rs

Srdja Janković
srdja.jankovic@udk.bg.ac.rs

¹ University Children’s Hospital, Tiršova 10, 11000 Belgrade, Serbia

² Astronomical Observatory of Belgrade, Volgina 7, 11000 Belgrade, Serbia

³ Future of Humanity Institute, Faculty of Philosophy, University of Oxford, Suite 8, Littlegate House, 16/17 St Ebbe’s Street, Oxford OX1 1PT, UK

biospheres, of a certain evolvability, may itself be regarded as a unit (indeed the arch-unit) of natural selection.

Keywords Evolvability · Natural selection · Astrobiology · Abiogenesis · Continuity thesis

Introduction

Laws of Nature are, as noted by Davies (2011) and many others, “blind to life”, not favoring in any way living states of matter over non-living states, much less making them a goal of cosmic evolution. At the very least, to assume otherwise would mean recourse to vitalism or worse. However, once life is established, it includes information carriers and executor mechanisms that are selected for increasing, or at least preserving, the likelihood of their own persistence. We do not yet know how exactly these information-carrying molecular patterns came to exist on Earth, much less how to generalize our consideration of life’s possible origins to patterns of any physical nature in any permissive environment, including exotic ones like subglacial oceans of Jovian moons, hydrocarbon lakes on Titanoid bodies, or Hoylean interstellar gas clouds. We do, however, know that for life to emerge there must be a qualitative shift from “pure” physical processes and chemical reactions, to such that encode and transfer specific types of information. This may be achieved by a particular kind of copying mechanism—one that copies the growth process of a prebiotic *ensemble*, rather than merely copying the product of this process (Woolf 2015). It may, therefore, be more than mere speculation if we attempt, even at the present unsatisfactory state of our knowledge about actual prebiotic systems, to create a draft representation of the processes of conception of life’s code(s) and encoding mechanisms. In doing so, we are essentially attempting to conceptualize the emergence of concepts.

Understanding Evolvability

The concept of evolvability is certainly not new. This concept was defined, accepted and applied in informatics long before it started to bore its way into biology, although the same concept is implicitly present in some proposed evolutionary mechanisms (e.g., genetic assimilation; Waddington 1953; Schmalhausen 1986). It has been formalized by Altenberg (1994) and applied to many evolving algorithms. Generalization of this concept to living things has been repeatedly attempted and fervently disputed. It has, however, seldom been explicitly included in analytical work on possible scenarios of biogenesis. Dawkins (1989) has argued that evolvability may evolve by a process of second-order selection. An overview of major implications of evolvable evolvability has recently been offered by Kirschner (2013). Evolvability might play a key role in extending the explanatory projects in physics based on the parameter spaces of possibility, such as Deutsch’s (2013) “constructor theory”, to living systems. A plausible role of evolvability and its origins in the extended synthesis of evolutionary theory has been discussed by Pigliucci (2008).

Evolvability of a biosphere is the measure of summary potential of evolutionary change of all its living beings, together with some measure of overall propensity of its systems to undergo evolutionary change upon given conditions. The analysis of evolvability teaches us that variations that are the substrate of evolution, though certainly “random” in the sense of not

being directed toward any pre-conceived goal, are nonetheless far from being mere permutations within the genetic script. Instead, they are changes of differing depth, operationally constrained by pre-existing regulatory networks, feedback loops, general genome structure and its plasticity, and multiple levels of selection. Therefore, in order to be able to fully understand evolution of a biosphere as an interconnected network of complex processes, it is necessary to account for the way evolvability itself is being altered by evolutionary change. Thus, in addition to evo-devo and eco-evo subfields of the extended modern synthesis of evolutionary theory, there is also need to explicitly include the “evo-evo” analytic approach.

Let us take the much honored example of Darwin’s finches of the Galapagos islands to briefly illustrate the above point. If we were to think that different selection pressures present on different islands of the archipelago were to have directly molded all the individual genetic elements required to alter the shape of the birds’ beaks in respective manner, it would be (as it was, indeed, often pointed out by critics of evolution) truly difficult to reconcile the observed effect with evolving subpopulation sizes, even on a quite long timescale. The seeming paradox is fully resolved by the fact that beak development consists of a number of intertwined, coordinated and co-regulated processes, this co-regulation being amenable to evolutionary change—in the case of the finches by alterations of the *Bmp4* gene, as shown by Abzhanov et al. (2004) and recently confirmed, to a great extent, by whole-genome analysis (Palmer and Kronforst 2015). One can see that the ability of the finches to evolve their respective beak shapes was dependent on the prior evolution of a genetic “handle” that allowed the relevant phenotypic characteristics to be accessible to evolutionary change without the requirement that each separate characteristic (or each of the involved genes) evolve independently.

Another striking example of evolved genetic architecture that allows a highly complex phenotypic pattern to evolve by mutation(s) of a single gene is the gender-restricted mimicry supergene *doublesex*. The function of this supergene was recently elaborated in butterflies of the genus *Papilio* (Kunte et al. 2014) and the possibility that this supergene plays a key role in mimicry in many species is currently being explored. Indeed, one should not be surprised if examples of analogous genetic mechanisms underlying rapid evolution of complex traits, in manner similar to mimicry wing patterns, begin to be discovered at a rapid pace throughout the tree of life.

Now, for such a “handle” to have evolved, there had to have been a prior evolution of a prerequisite genetic/regulatory super-structure. We might name the basic unit of this structure the *evolvon* and tentatively define this term as an entity that consists of (i) a node within genetic script where regulatory networks intersect in such a manner that a mutation may produce a coordinated change in multiple traits and functions; together with (ii) all genes with expression critically influenced by the node. The development of such a super-structure in a realistic time-frame cannot, however, be fully explained as a simple result of natural selection acting directly and independently upon its components and internal relationships. One must instead assume that the selection for appropriately evolvable systems is the aforementioned higher-order selection, which needs to be understood as strongly time-dependent, in the sense that its effects are only seen *a posteriori* by tracing back the conceptual qualities of taxa (or biospheres) that are extant at the time-point of analysis. Even though we do not have empirical access to more than one coding concept, this view of selection is not necessarily contrary to the Williams principle, that postulates selection *between* entities at a given level of biological organization as a requirement for evolutionary adaptation at that level. The proposed idea is, however, clearly opposed to the Maynard-Smith principle, that demands that there be no selection *within* entities of the considered level (for an excellent discussion of implications of these principles, please see Gardner 2014).

In our examples, there was a compound process selecting the total posterity of the ancestral finch population that flew over to the Galapagos or the total posterity of the ancestral butterflies with appropriately evolvable wing patterns: in the case of the former, selecting for the ability of the finches to survive in different habitats by diversifying under differing selection pressures; in the case of the latter, selecting for those butterflies most likely to evolve successful mimicry capabilities. In this second-order selection analysis, the unit of selection is the whole sum of potential populations that can radiate from the ancestral one, while the selected trait is, in effect, the section of morphospace accessible to this ancestral population on the basis of its genetic structural/functional disposition – what we might choose to name the “*evolvome*”.

The evolution of evolvability, i.e., selection for genetic alterations conferring greater likelihood of acquiring evolutionary change in some characteristic(s) of the organism, has also been incontrovertibly observed in real time, at least for the finest branchings of trajectories within evolvability space, such as evolutionary adaptations arising in the setting of pathogen-host interactions. One of many documented examples is the ability of bacterium *Borrelia burgdorferi* to evolve additional evolvability towards acquiring alterations of key antigen epitopes recognized by its host's immune system—alterations that confer clear survival advantage to the organism by enabling it to escape immune responses (Graves et al. 2013). The effect of coevolution on evolvability has also been explored and confirmed by mathematical modelling (“evolutionary experiment”) performed by Zaman et al. (2014). Of course, such possibilities are still underexplored in evolution studies, since the experimental timescales required for manifestations of differential evolvability are usually impractically large for all organisms except those with very brief generation times (i.e., microorganisms). Therefore, researchers are most often compelled to use indirect (e.g., paleontological) evidence.

Coding Concept(s)

All currently observed life on Earth shares a common origin. This fact, supported by vast amount of evidence, is reflected in the information-encoding and information-processing systems that are accordingly shared by all known life-forms that inhabit this planet. This semantics of life, embodied by nucleic acids that store information in the sequence of nucleotide triplets (“letters”) and the syntax embodied in the specific relationship between the genetic code and amino acid sequences of proteins, together with the grammar conferred by regulatory networks, constitute the language of life. This language is supported by the physical-chemical substrate that allows “top down” as well as “bottom up” flow of information, as discussed by Walker and Davies (2013). This flow is the basis of hierarchical structure of living systems. Of note, in this context, information is causally efficacious and plays an active role in controlling the properties and behaviour of the system, maintaining its homeostasis, and, ultimately, determining its ability to survive and reproduce under given conditions. This causal role of information (algorithmicity or algorithmic takeover) that is unique to living systems has, indeed, been proposed as a key to our quest to define what life is (Walker and Davies 2013). Accepting such an informatic definition of life was actually a necessary requirement for the present work.

For the purpose of this preliminary (or, indeed, pre-preliminary) discussion, we shall provisionally define the “coding concept” as the sum of distinct stable (as well as transferable) physical/chemical system states that are available as a substrate for holding heritable

information, together with the “rulebook of meanings” that governs the flow of this information and its authority in shaping, controlling and regulating the system (and, at the same time, being shaped, controlled and regulated by the system). In the only empirically accessible example so far, our earthly nucleic-acid based life, the chemical substrate is supplied by four¹ different nucleotides, while the “coding language” of their triplets (codons), and the way they govern protein synthesis, provides the rulebook.

Though evolution of life on Earth has led to immense diversity of structures and mechanisms used by living beings to ensure their survival and propagation, it is evident that the core attributes of the coding concept (language of life) have changed very little, if at all, since the putative Last Universal Common Ancestor (LUCA) and almost certainly a considerable amount of time before. Indeed, the meaning of codons, as well as the way key information stored in DNA is used to enable life’s processes, is (almost) exactly the same for a bacterium and a blue whale. If we go back to the beginnings of life, whether we seek them in Darwin’s “warm little pond” filled to the brim with Urey-Millerian prebiotic chemicals, or in far more exotic places, it becomes clear that life’s coding concept(s) that had originally evolved, be there one or many, have been effectively tested by the whole extent of history of life on Earth. And – since there is currently only one empirically accessible coding concept – one biosphere of closely related beings—it is legitimate to view the whole of life on Earth as a unique result of this selection: a single evolved entity.² Whether or not there had also been other, failed coding concepts that did not pass selection, the one that we now observe can be said to have been selected by its ability to give rise to viable life all the way up to the present—selected for its survivability during at least 3.4 (and probably closer to 3.8) billion years. Terran life has persisted and repeatedly risen to the challenge of surviving mass extinctions caused by dramatic changes in living conditions on the planet through this time interval. The total potential of life to adapt, endure and diversify is, when observed from the “zoomed out” four-dimensional spatiotemporal point of view, a function of its coding concept. This is true because, in order for life to be able to evolve through a combination of random mutations and (non-random) natural selection, the coding concept must provide a balance between stability of hereditary information and ability to open avenues of change. The dominant role of the coding concept in offering a smörgåsbord of possibilities for life’s adaptation to challenges is also reinforced by various forms and mechanisms of lateral gene transfer, particularly important in the early development of the terran biosphere (Vetsigian et al. 2006).

Each coding concept determines a particular subspace of the total parameter space of possible life-forms. It also provides an *evolvability landscape*, i.e., differential likelihood of changes leading toward any particular adaptation (along an evolutionary track). Natural selection determines³ whether a particular (open) track will be travelled by a particular population at a particular time; the existence and extension of tracks themselves is, however, predetermined by the characteristics (content) of information present in the evolving population at the time we choose to regard as the starting point (or, more properly, branching point; for a comprehensive analysis of branching [bifurcation] in evolutionary trajectories from the perspective of non-linear dynamics, please see Volkenstein 1987). This analysis of constraints (evolvability) vs. selection pressures is integral part of the classic (and widely accepted)

¹ There are, of course, five different nucleotides in biology, but no individual nucleic acid present in living systems uses all five, therefore the semantics of life is based on four.

² This must not be confused with the Gaia hypothesis in its stronger forms (ascribing to this entity key attributes of an “organism” and even self-consciousness), although it is clearly not incompatible with it.

³ Or rather sub-determines, given that there are also other processes at play (e.g., genetic drift).

representation of processes whereby new “larger” or “smaller” taxa originate: the larger the section of the parameter space that is “opened up”, or made accessible, by a particular change, the greater the taxon-building potential. However, the ability of a given genome alteration to open the access to a particular tract of the morphospace is, in addition to the specific nature of the alteration in question, heavily dependent on core characteristics of the previously built systemic *context* that allows the alteration to have a far-reaching effect on ability to evolve. This ultimately goes all the way to the general principles of organization of living systems, such as modularity, pleiotropy, and redundancy—principles that command that evolvability be inevitable (Lehman and Stanley 2013).

Here, we chose to leave out discussion of phenotypic plasticity, together with all issues of genotype-phenotype relationships that certainly have the potential to render the analysis much more complex. We felt free to do so because we hold that, even with a great area of freedom for the phenotype, it is still, at least to a significant extent, constrained (sub-determined) by genotype, conferring a less-than-random relationship between coding space and morphospace. This is not to gainsay that plasticity at all levels takes important part in the overall process of evolution and certainly needs to be addressed in future work.

The key question is: is it possible to regard the core coding concept of information processing in a particular type of living system as the primordial unit of selection? The establishment of this core coding concept must be seen as the crucial event that (1) “opened up” the parameter space that includes all extant and extinct living beings on Earth, as well as the vast, but finite *ensemble* of all living beings that could exist, or could have existed, based on the same coding concept (i.e., DNA/RNA/proteins etc, with existing rules for information storage and usage within this substrate) and (2) provided the “toolbox of toolboxes”, containing the primary tools to build all the tools (modules, regulators, checkpoints, etc.) needed to evolve the evolvability that is prerequisite for life to persist and diversify. In other words, we entertain the question whether the coding concept of “our life” could itself be a result of natural selection—selection resulting in today’s single entity manifested by the whole biosphere of Earth. As aptly demonstrated by Doolittle (2014), the very fact that this biosphere has survived to the present can be taken to imply a form of natural selection operating on alterations, accumulated in time, that enhance global survivability (i.e., mechanisms that allow novel ways of coping with drastic changes in planetary conditions, such as the “snowball Earth” episodes or the Chicxulub impact). It is, in turn, logical to expect and relatively easy to demonstrate that survivability, particularly at the biospheric level, will tend to be in a significant correlation with evolvability.

Walking Down “Evolvability Alley”

If we wish to hold any teleological explanations at bay, there is no alternative explanation for the emergence of evolvability than that it had itself had to have evolved. Since we see the outcome of the collective processes of selection, but not the original sample of selection units (taking the coding concept of “this life”, and the biosphere that it achieved, as the selection unit), we may regard the general case (one that, at least by logic, may be expected to apply to any habitable⁴ planet) as a collection of prebiotic substrates, that vary within given physical

⁴ We may use a more or a less restrictive definition of habitability, depending on whether we wish to include more “exotic” possibilities of life/coding concepts; this does not change the point.

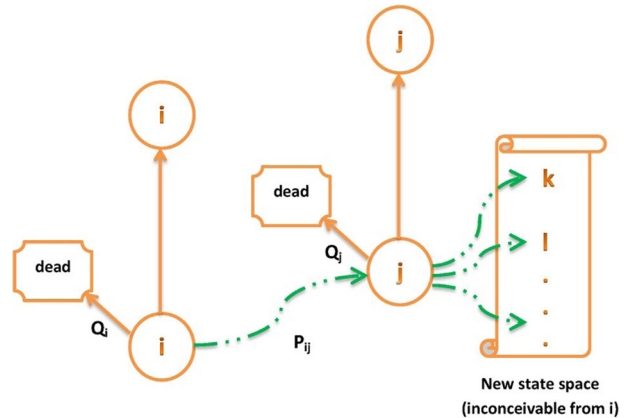
constraints subject to the continuity thesis of Fry (1995, 2000), and give rise to a number of different “coding concepts” with *differential evolvability*. These may differ along more than one dimension: for instance, one might imagine life involving DNA, but with different meaning of codons; on the other hand, one might also think of life using different nucleic acids, no nucleic acids at all, no carbon-based complex molecules (“organics”), or even no complex *chemical* patterns but some other (though, emphatically, still physical) information carrier(s). The more evolvable ones are evidently more likely to give rise to survivable biospheres; however, at some theoretical limit, the tug-of-war between evolvability and stability of heritable information must surely be mutually balanced at the scales of survivability, defining the other boundary of the “evolvability alley”—the path through the evolvability parameter space that gives sufficient evolvability without critical degeneration of information transfer. Needless to say, this evolvability alley is itself multidimensional, since for every conceivable direction within the morphospace/coding space there is an associated evolvability measure.

A schematic diagram of a very simplified model of such evolvability effects in the context of biospheric evolution is shown in Fig. 1. Transition through the evolvability alley has been represented by green dash-dot arrows, in contrast to “usual” processes such as persistence and extinction. The major difficulty in any such presentation is that opening of completely new state spaces in the high-evolvability regime means adding new dimension to morphospace—essentially that one needs new metric at each particular step. This obviously leads to huge practical difficulties in building quantitative models of the whole process. High evolvability creates a sort of horizon of (un)predictability surrounding each of the key evolutionary steps. If some biospheres in the entire ensemble pass from the state i into state j , with a finite probability P_{ij} , new structures added in that process are likely to open a series of new target states of future transitions (k, l, \dots), which cannot be envisioned on the basis of even a very deep analysis of the state i . It is to be expected, however, that the probabilities of extinction will change, possibly dramatically, even if the physical causal mechanisms stay the same during the entire biospheric history. Roughly, Fig. 1 is complementary to Fig. 3 in Doolittle (2014), offering further intuition into why we can reasonably consider biospheric selection as a Darwinian process. We shall report some preliminary results of numerical models of this type, representing the emergence of key evolutionary innovations within a particular biosphere as a Poissonian process, in a forthcoming work. For the present, we emphasize how it comes entirely naturally to speak about capacity of natural selection to create *effective* adaptations even to the largest, fauna-changing cataclysms; not at the level of populations, species and higher taxa, but at the level of biospheres themselves. The relevant structures in morphospace will be detectable, of course, only *post festum* in analysis of the ecological properties of the survivors. While this is clearly an unorthodox way of referring to the power of natural selection, there appears to be no reason to consider it any less analytically rigorous.

The “survivability-enhancing mutations” mentioned by Doolittle, in the above perspective, become rather events that open up new dimensions of evolvability. An appropriate metaphor for this is, perhaps, the Tower of Babel, as envisioned by Pieter Bruegel the Elder (cca. 1563, especially the small or Rotterdam version; Fig. 2)—each subsequent spiral tier standing on the preceding tier and sharing with the latter a common basic structure—the “coding concept”.⁵ It is clear that the survivability (successful persistence) of the highest tier actually *selects* all the

⁵ This metaphor is also appealing because of the Biblical ending of the narrative of this project – the Tower ceases to grow when the “coding concept” becomes critically degenerated (there is no longer a commonly understood language).

Fig. 1 A symbolic flowchart of biospheric evolution



tiers below. If we stand at the N^{th} tier and look down, we may perceive tiers all the way to the ground—but, due to the anthropic shadow (Ćirković et al. 2010), we can only find such that are inherently capable of supporting those tiers that allowed us to have come to be. Crucially, since we see (at this time) no other towers but ours, we cannot know how many of them—in various stages of growth – there are or how many there may be.⁶

Therefore, results of natural selection perceived on the basis of its objective outcome (biosphere of Earth as seen to exist) must not be confused with artifacts of an observer selection effect – the fact that it is *us* who see the Earth’s biosphere now and *we* cannot reasonably expect to see anything incompatible, or unlikely to be compatible, with the processes necessary for *our* emergence and survival to the present. This is the anthropic bias, often wrongly teleologically interpreted (cf. Ćirković et al. 2010). The situation in which there are no observers evolved to argue for a different type of biogenesis and subsequent evolution might logically occur elsewhere, in a locale described by different sets of lower-level parameters: “wrong” type of planet, “wrong” position in a galaxy, even “wrong” set of effective low-energy physical laws. However, this observation selection effect does not diminish the amount of explanatory work we need if we are to account for those properties.

The “Coding-Conceptual Space”

Let us borrow a metaphor from chess: after a move has been drawn, the space of all possible moves is sectioned into (i) those that are still possible in subsequent steps, and (ii) those that are no longer available as an option for the player. Those subspaces are exactly measurable for each move made in each given position, although we might not have an algorithm for predicting these measures in advance. We may say that events giving rise to new taxa are like chess moves (the other side of the chessboard, in this case, is not occupied by an adversarial player, but may serve as a sub-metaphor for existential threats to survival⁷). We may define as

⁶ This still prompts us to ask questions about specific construction processes used at each tier – to continue with our fine-arts metaphor; we can do so with Klein (1978).

⁷ It is tempting to choose black for this side, but in the homeland of chess it would almost certainly be more naturally regarded as white, since, contrary to most Europeans, Indian cultures tend to associate death with white rather than black.

Fig. 2 The Tower of Babel, cca. 1563. (Courtesy of the Google Art Project at <https://www.google.com/culturalinstitute/project/art-project>.)



successful those moves that allow us to proceed further along the course of the game, disregarding the vain hope that we may win by checkmate, as impossible for us as it was for Antonius Block (Bergman 1957). The further we reach, the finer are the subdivisions of the parameter space of remaining possible moves, until the breadth of space of available decisions and their outcomes becomes reduced to that of a typical recreational chess problem, solvable by simply visiting all possibilities.

Let us now return to our chessboard in the instant before the first move. We may look upon the total parameter space allowed by the chessboard layout and the rules of the game (all possible sequences of moves in all possible games) as a single entity. It is defined by the “coding concept” of the game. And, reminiscent of a well-known poem of Jorge Luis Borges, we may well ask ourselves if there could be something *beyond* the choice of this particular coding concept that defines this particular game we call chess – amid the parameter space of all possible games, with same or widely different board, figures and rules.

The chess metaphor may be brought closer to an actual analogy if we pause to consider that chess is actually selected by human beings for its “playability”, the latter being dependent both on intrinsic features of the game and extrinsic whims of cultural evolution, which is testified by the fact that standard chess has proved to be exceedingly stable against various modifications (“chess variants”), even those proposed and promoted by luminaries such as Capablanca, Fischer or Yoko Ono. However, the chess metaphor is here clearly more powerful than the chess analogy, leading us to regard the “Library of Mendel” imagined by Daniel Dennett (Dennett 1995) as a single volume somewhere in the N-dimensional “Library of Libraries”, containing not merely all possible books written in an alphabet, but a collection of collections of all possible books written in all possible alphabets—both in terms of letters/symbols and reading modes. And, inevitably, we ask ourselves—what makes some alphabets more “literaturable” than others? It appears reasonable to accept a rather mundane idea that the potential of a text to clearly, enduringly and efficiently convey information in a multitude of diverse and unpredicted contexts is, ultimately, a function of the coding concept at the core of each alphabet. Such an assumption is analogous to our hypothesis that the capacity of life to adapt, endure and diversify is, ultimately, a function of its coding concept within each biosphere or a class of biospheres.

Implications for Astrobiology

What are the possible applications of the above reasoning useful to the astrobiological quest and the understanding of life in general? While it may be impossible to generalize, even in Bayesian terms, on universal likelihood of biogenesis based on Earth's lone example (Spiegel and Turner 2012), this "sample of one" "life as we know it" is, in fact, the result of operation of long-acting selection processes, selecting (or at the very least incessantly re-affirming) the coding concept that this life, and this particular biosphere, is built on. It is plausible to assume that many possible avenues of prebiotic systems' evolution are almost certain to provide a number of different coding concepts. In theoretical treatment of shadow biospheres (Davies et al. 2009; Davies 2011), the prevailing explanation for their absence (or rather our failure to detect one so far) is that life belonging to the visible biosphere, with its successful coding concept, simply outcompeted any possible "others" for resources. However, in accordance with Doolittle's analysis (2014), such competition, or indeed any form of direct or indirect antagonism, or even any interaction at all, may not be an absolutely necessary part of the explanation, since differential survivability is, in principle, sufficient and compatible with what we observe.

Explicit addition of the "evolved evolvability" concept to the current framework of understanding the processes that dictate the emergence of life, on Earth or anywhere else, results in what might be called the "expanded continuity thesis"—the notion that life's evolvability is an evolved ability – an ability that is actively selected for by the whole system wherein life originates and in which it becomes embedded. This system includes both the home planet (the "cradle") and its environment, potentially encompassing the whole galaxy, or even the whole cosmological domain, for it is by now quite proven that life's history (and thus any occurring survivability selection) is heavily influenced by factors originating outside of, and potentially many parsecs away from, the home planet. The fact that habitable planets are *open systems*, and that therefore their evolution is coupled to wider astrophysical processes, is yet another indication of the disteleological nature of universal evolutionary process. To claim a significant likelihood that any given sample of prebiotic systems, if sufficiently large and endowed with enough variability, will contain at least one that is sufficiently evolvable to give rise to a successful biosphere (in terms of long-term survival), effectively constitutes another way to state the continuity thesis. This is also in accordance with the NASA definition of life (Des Marais et al. 2008).

There is another reason why enhanced evolvability might provide decisive advantage to those biospheres within which it emerges, which is in complete accordance with the prevailing "neocatastrophic" views in astrophysics, planetary and geo-sciences. Let us suppose that there is a series of *temporal windows* for achieving any crucial evolutionary transition, the boundaries of which are defined by global upheavals, such as occasional asteroidal/cometary impacts, supervolcanic episodes, or nearby supernovae/ γ -ray bursts. Then, there will be a selection pressure among biospheres for evolving at least one taxon (per potential threat) capable of surviving and transmitting the innovation into the next window of opportunity. Obviously, higher evolvability will lead to increased probability of evolving such a taxon in a fixed amount of time (very weak assumptions about stochasticity of evolutionary mechanisms may suffice for this conclusion). We might not be able to infer this directly, however, because a sort of inverted Signor-Lipps effect will act to artificially suppress the apparent diversity of the survivors. We cannot delve here into more details on this aspect of the biospheric selection (see for instance Ćirković and Vukotić 2008 and references therein for some of the astrobiological ramifications).

The concept of evolved evolvability, as envisaged above, apart from other implications, perfectly counterbalances “lucky accident” arguments for rarity of life in the Universe, based on the perception that Earth life is dependent on extremely fine tuning (Barnes 2012) of a mind-bogglingly vast number of parameters at all levels. In a way, evolved evolvability turns the table on this school of thought by arguing that Life, if it takes hold at all, *necessarily* comes with a built-in all-purpose fine-tuner mechanism. Indeed, one would not err to say that Life *is* this fine-tuner. And the operating principle of the tuner, in essence, consists of well-known Darwinian/Wallacean forces of natural selection acting upon coding concepts as selection units. The concept of coding concept as the ultimate selection unit of life is a natural extension and, in a way, completion of analysis of the hierarchy of natural selection, an important question in general evolutionary theory, masterfully reviewed by Stephen Jay Gould in his essay “Guliver’s Further Travels” (Gould 1998). This is also, in a certain sense, an extension of widely accepted evolutionary thinking style that “if there is a solution, life will find it”—to the effect that “if there is a coding concept for life (hidden somewhere in the conceptual space) that may work (i.e., become sufficiently evolvable to be surviving of challenges) in this particular spatiotemporal location, it will eventually be realized and selected”. The whole downstream history from the point of origin (taken as any form of local abiogenesis, but optionally extended to planetary seeding in various scenarios of panspermia) is here seen as a unique instrument of selection that retains or eliminates a coding concept, as it retains or eliminates a biosphere that has built itself out of one.

Conclusion

The most important consequence of the above analysis for astrobiological research might be the need to supplement the classical concept of habitability with some measure of “physical conceivability”, defined by the existence of physical conditions that allow (or, indeed, dictate) the emergence of (physical/chemical) information carriers capable of supporting coding concepts that may then evolve towards evolvability and eventually be selected for their respective biosphere—building potentials. Although the hypothesis that we put forward is notoriously difficult to test empirically, potential avenues of research contributing to this goal clearly exist. Apart from the continuing quest for examples of life outside Earth, investigations of possible shadow biospheres hidden on our planet (Davies 2011), and diverse mathematical models of evolvability under different conditions and constraints, informed by extrapolations of present physical knowledge and/or information-theoretical considerations (Adami 2015), the phenomenon of evolvable evolvability by natural selection at the level of the coding concept could also be experimentally tested in biological laboratories, by looking at evolvability of differently conceived life-forms. This may be accomplished imminently, using, for example, semi-synthetic organisms with different genetic alphabet (Malyshev et al. 2014; Dhami et al. 2014; Cleaves et al. 2015) or, alternatively, semi-synthetic (or even fully synthetic) organisms with differentially designed systems for protein synthesis (Daube and Bar-Ziv 2013). Of note, the former could be a tool to test the evolvability-associated effects of different semantics, while the latter would enable us to test variations of syntax. Incorporation of either or both into more complex systems could, in principle, provide a way to investigate the grammar as well. *In vitro* evolution experiments also offer the promise to enable investigations into contextual dependence of a given coding concept, beginning, naturally, with our own (Popović et al. 2015).

All of this invokes a new area of research, the aim of which is to define parts of the “conceptual space” that are accessible under given physical conditions and plausible to be accessed on a given timescale. It is clearly a task for interdisciplinary teams comprised of theoretical physicists, astrophysicists, Earth/planetary scientists, information theorists, computational chemists, evolutionary biologists, systems biologists, biosemantic experts, comparative linguists and researchers from other fields, as well as philosophers. Ideally, the results of such endeavor might allow us to gain an improved understanding of what we do when we devise astrobiological research strategies and, hopefully, to tailor the design of our future life-seeking expeditions, whether they take place on Earth, in the Solar System or beyond, to our ability to grasp the very essence of coding concepts as the primordial selection units of Life.

Acknowledgments Our foremost thanks go to Eva Kamerer, Ana Parabucki, Sergej Tomić, and the two anonymous reviewers for *Origins of Life and Evolution of Biospheres* for their thoughtful comments and pertinent criticisms of a previous version of this manuscript. We have also benefited from insightful discussions of related topics with Jelena Dimitrijević, Aleksandar Obradović, Jovana Kabić, Tanja Berić, Kim Sterelny, Slobodan Perović, Dušan Pavlović, Anders Sandberg, and Branislav Vukotić. This research has made use of NASA Astrophysics Data System and the KoBSON consortium of Serbian libraries. The authors have been partially supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia through grants ON41004 (S. J.) and ON176021 (M.M.Ć.).

References

- Abzhanov A, Protas M, Grant BR, Grant PR, Tabin CJ (2004) Bmp4 and morphological variation of beaks in Darwin’s finches. *Science* 305:1462–1465
- Adami C (2015) Information-theoretic considerations concerning the origins of life. *Orig Life Evol Biosph* 45: 309–317
- Altenberg L (1994) The evolution of evolvability in genetic programming. In: Kinnear KE Jr (ed) *Advances in genetic programming*. MIT Press, Cambridge, pp 47–74
- Barnes LA (2012) The fine-tuning of the universe for intelligent life. *Publ Astron Soc Aust* 29:529–564
- Bergman I. *Det sjunde inseglet*. Svensk Filmindustri 1957
- Ćirković MM, Vukotić B (2008) Astrobiological phase transition: towards resolution of Fermi’s paradox. *Orig Life Evol Biosph* 38:535–547
- Ćirković MM, Sandberg A, Bostrom N (2010) Anthropocentric shadow: observation selection effects and human extinction risks. *Risk Anal* 30(10):1495–1506
- Cleaves HJ II, Meringer M, Goodwin J (2015) 227 views of RNA: is RNA unique in its chemical isomer space? *Astrobiology* 15(7):538–558
- Daube SS, Bar-Ziv RH (2013) Protein nanomachines assembly modes: cell-free expression and biochip perspectives. *WIREs Nanomed Nanobiotechnol* 5:613–628
- Davies PCW (2011) Searching for a shadow biosphere on Earth as a test of the “cosmic imperative”. *Phil Trans R Soc* 369:624–632
- Davies PCW, Benner SA, Cleland CE, Lineweaver CH, McKay CP, Wolfe-Simon F (2009) Signatures of a shadow biosphere. *Astrobiology* 9:241–249
- Dawkins R. The evolution of evolvability. In: C Langton (ed). *Artificial Life*. Addison-Wesley 1989; pp. 201–20.
- Dennett CD (1995) Darwin’s dangerous idea: evolution and the meanings of life. Simon & Schuster, New York
- Des Marais DJ, Nuth JA III, Allamandola LJ, Boss AP, Farmer JD, Hoehler TM, Jakosky BM, Meadows VS, Pohorille A, Runnegar B, Spormann AM (2008) The NASA astrobiology roadmap. *Astrobiology* 8:715–730
- Deutsch D (2013) Constructor theory. *Synthese* 190:4331–4359
- Dhami K, Malyshev DA, Ordoukhanian P, Kubelka T, Hocek M, Romesberg FE (2014) Systematic exploration of a class of hydrophobic unnatural base pairs yields multiple new candidates for the expansion of genetic alphabet. *Nucleic Acids Res* 42(16):10235–10244
- Doolittle WF (2014) Natural selection through survival alone, and the possibility of Gaia. *Biol Philos* 29(3):415–423

- Fry I (1995) Are the different hypotheses on the emergence of life as different as they seem? *Biol Philos* 10(4): 389–417
- Fry I (2000) *The emergence of life on Earth: a historical and scientific overview*. Rutgers University Press, New Brunswick
- Gardner A (2014) Life, the universe and everything. *Biol Philos* 29:207–215
- Gould SJ (1998) Gulliver's further travels: the necessity and difficulty of a hierarchical theory of selection. *Philos Trans R Soc Lond B Biol Sci* 353(1366):307–314
- Graves CJ, Ros VID, Stevenson B, Sniegowski PD, Brisson D (2013) Natural selection promotes antigenic evolvability. *PLoS Pathog* 9(11):e1003766
- Kirschner M (2013) Beyond Darwin: evolvability and the generation of novelty. *BMC Biol* 11:110
- Klein HA (1978) Pieter Bruegel the elder as a guide to 16th century technology. *Sci Am* 238(3):134–140
- Kunte K, Zhang W, Tenger-Trolander A, Palmer DH, Martin A, Reed RD, Mullen SP, Kronforst MR (2014) *Doublesex* is a mimicry supergene. *Nature* 507:229–232
- Lehman J, Stanley KO (2013) Evolvability is inevitable: increasing evolvability without the pressure to adapt. *PLoS One* 8(4):e62186
- Malyshev DA, Dhami K, Lavergne T, Chen T, Dai N, Foster JM, Corrêa IR Jr, Romesberg FE (2014) A semi-synthetic organism with an expanded genetic alphabet. *Nature* 509(7500):385–388
- Palmer DH, Kronforst MR (2015) Divergence and gene flow among Darwin's finches: a genome-wide view of adaptive radiation driven by interspecies allele sharing. *Bioassays* 37(9):968–974
- Pigliucci M (2008) Is evolvability evolvable? *Nat Rev Genet* 9(1):75–82
- Popović M, Fliss PS, Ditzler MA (2015) *In vitro* evolution of distinct self-cleaving ribozymes in diverse environments. *Nucleic Acids Res* 43(14):7070–7082
- Schmalhausen I (1986) *Factors of evolution: the theory of stabilizing selection*. University of Chicago Press, Chicago
- Spiegel DS, Turner EL (2012) Bayesian analysis of the astrobiological implications of life's early emergence on Earth. *PNAS* 109(2):395–400
- Vetsigian K, Woese C, Goldenfeld N (2006) Collective evolution and the genetic code. *Proc Natl Acad Sci U S A* 103(28):10696–10701
- Volkenstein MV (1987) Punctualism, non-adaptationism, neutralism and evolution. *BioSystems* 20:289–304
- Waddington CH (1953) Genetic assimilation of an acquired character. *Evolution* 7:118–126
- Walker SI, Davies PCW (2013) The algorithmic origins of life. *J R Soc Interface* 10:20120869
- Woolf NJ (2015) A hypothesis about the origin of biology. *Orig Life Evol Biosph* 45:257–274
- Zaman L, Meyer JR, Davengam S, Bryson DM, Lenski RE, Ofria C (2014) Coevolution drives the emergence of complex traits and promotes evolvability. *PLoS Biol* 12(12):e1002023