Physical foundations of biological complexity

Yuri I. Wolf[a], Mikhail I. Katsnelson[b], and Eugene V. Koonin[a,1]

[a]National Center for Biotechnology Information, National Library of Medicine, Bethesda, MD 20894; and [b]Institute for Molecules and Materials, Radboud University, 6525AJ Nijmegen, The Netherlands

Biological systems reach hierarchical complexity that has no counterpart outside the realm of biology. Undoubtedly, biological entities obey the fundamental physical laws. Can today's physics provide an explanatory framework for understanding the evolution of biological complexity? We argue that the physical foundation for understanding the origin and evolution of complexity can be gleaned at the interface between the theory of frustrated states resulting in pattern formation in glass-like media and the theory of self-organized criticality (SOC). On the one hand, SOC has been shown to emerge in spin-glass systems of high dimensionality. On the other hand, SOC is often viewed as the most appropriate physical description of evolutionary transitions in biology. We unify these two faces of SOC by showing that emergence of complex features in biological evolution typically, if not always, is triggered by frustration that is caused by competing interactions at different organizational levels. Such competing interactions lead to SOC, which represents the optimal conditions for the emergence of complexity. Competing interactions and frustrated states permeate biology at all organizational levels and are tightly linked to the ubiquitous competition for limiting resources. This perspective extends from the comparatively simple phenomena occurring in glasses to large-scale events of biological evolution, such as major evolutionary transitions. Frustration caused by competing interactions in multidimensional systems could be the general driving force behind the emergence of complexity, within and beyond the domain of biology.

evolution of complexity | competing interactions | frustrated states | spin glasses | self-organized criticality

An oft-repeated quip attributed to Ernest Rutherford goes "all science is either physics or stamp collection." This tends to annoy scientists to no end due to the flagrant disregard of nonphysical science that it espouses. However, if one looks beyond the purported insult, the phrase seems to have a serious, even fundamental, meaning. Indeed, it emphasizes the major distinction between those scientific endeavors that strive to explain reality in terms of universal, microscopically reversible laws and not concerned with particularities of the real-world history and those that depend on history in their explanatory frameworks, such as (most of) biology or geology. The first class of approaches can be collectively denoted physics (1). There is no corresponding single term for the second type of scientific endeavors, but for the sake of discussion, we can call them "historical sciences," to emphasize their intrinsic time-irreversibility and dependency on unique events. In historical sciences, universal laws figure only in the background (although the fundamental laws of physics certainly hold), predictability and reproducibility become problematic, and any generalization is suspect (biologists seem to rather delight in saying that "there are exceptions to everything in biology")—hence the "stamp collection" moniker, whether or not one views it as derogatory. However, as aptly pointed out by Smith and Morowitz (2), "biology has traditionally occupied a middle ground between the determinism of classical physics and the uncertainties of history." Hence, an important goal of biological theory is to explore the extent to which more physics can be brought into our understanding of biology in general and evolution in particular (3).

Currently, fundamental laws of physics are considered to be local in space and time. Beginning with Newton's *Principia* (4),

physicists have held that the fundamental laws are expressed in terms of ordinary differential equations, that is, the whole evolution of a (mechanical) system is uniquely determined by coordinates and velocities of all particles at a given time instant. In classical field theory, the fundamental laws are represented by partial differential equations, that is, they are local not only it time but also in space. Indeed, this is the only type of physical laws that are consistent with general relativity theory, which shuns instantaneous interaction at finite distances (5). Then, how is it possible that many physical systems have history (in other words, memory of events past), sometimes going back for as long as billions of years? What are the physical mechanisms that could be responsible for the long-term memory in such systems? Coming back to Rutherford, how does "stamp collection" emerge in the world of physics?

There are major research areas that traditionally fall within the domain of physics, but where history matters. This is the case for all nonergodic systems of which structural and spin glasses (6, 7) arguably are the best-studied class. In physics, ergodic systems are defined as those for which the time spent by the systems in a region of the phase space of microstates is proportional to the size of that region. In other words, in the long term, all accessible microstates are equiprobable and are all eventually visited, so that over a long time the state of the system is independent of the initial state. In contrast, nonergodic systems do not visit all microstates and possess memory of the initial state (8, 9).

Glass-like systems are non-Markovian, that is, are characterized by historical memory whereby the present state of the system depends on the entire previous history, and accordingly the future states are not precisely predictable. This nonergodic behavior is caused by competing short- and long-range interactions which result in frustration and can produce complex patterns (10). Notably, the glass-like phases share with biological systems at least two fundamental features: (*i*) historical memory and the

Significance

Living organisms are characterized by a degree of hierarchical complexity that appears to be inaccessible to even the most complex inanimate objects. Routes and patterns of the evolution of complexity are poorly understood. We propose a general conceptual framework for emergence of complexity through competing interactions and frustrated states similar to those that yield patterns in striped glasses and cause self-organized criticality. We show that biological evolution is replete with competing interactions and frustration that, in particular, drive major transitions in evolution. The key distinction between biological and nonbiological systems seems to be the existence of long-term digital memory and phenotype-to-genotype feedback in living matter.

resulting contingency (nonergodicity) and (*ii*) complexity. These parallels have been drawn previously by Laughlin and Pines (11) and Laughlin et al. (12), who suggested that modern physical theory of glasses might substantially inform different areas of theoretical biology. Here we develop the idea of the crucial, central role of competing interactions and frustration in biological evolution by examining specific biological concepts and phenomena and comparing their behavior to that of glass-like systems. In contrast with the previous analyses (11, 12), which primarily emphasized glass-like features of biological macromolecules (see also ref. 13), and in particular the role of frustrating interactions in protein folding (14, 15), we focus on frustrations and competing interactions in evolutionary dynamics.

Biological evolution undeniably leads to the emergence of complexity (16–22). Moreover, complexity has been proposed as an identifying feature of "artefacts," that is, systems of (ultimate) biological origin. In a (so far largely gedanken) search of putative extraterrestrial habitats of life, any objects of "unreasonable complexity" could be the principal signatures of biological activity (23). Complexity is notoriously difficult to define precisely, and no single, universal definition has been agreed upon (24–26), although it seems to be commonly held that "when we see it, we know it." The definitions of complexity that appear to be meaningful in biology involve, depending on the level of analysis, the number of evolutionarily conserved nucleotide sites in genomes and the number of genes or functional components in organisms or suborganismal functional systems, as well as the hierarchical organization of biological entities, be it functional networks and pathways, cells, organisms, or communities (20–23, 27). Perhaps the most general definition, pathway complexity, deriving from the concept of algorithmic complexity in mathematics, includes the number of steps required to create a given object. It has been proposed that entities with a pathway complexity above a certain threshold can only originate from biological processes (23).

The origins of biological complexity have been approached from both biological and physical standpoints. Traditional biological narratives view complex features as adaptations. However, more recently, neutral scenarios for evolution of complexity have been proposed (28–33). Under these models, complexity ensues from evolutionary accumulation of genomic features, such as duplicated genes, introns, or mobile genetic elements (MGE). Such features accumulate nonadaptively in organisms with small effective population size, in which purifying selection is too weak to purge them. The genomic embellishments that have not been discarded can gain function, for instance, under the subfunctionalization model of evolution of duplicated genes, when the duplicates differentially lose subfunctions of the ancestral genes. This nonadaptive process translates into "constructive neutral evolution" whereby complexity emerges in the form of interdependence of subfunctionalized components of an evolving cellular system (34–37). However, adaptive origin of complexity cannot be discarded either, as suggested, for example, by recent models of prokaryotic genome evolution indicating that acquired genes are, on average, beneficial, conceivably, thanks to the increased functional versatility of more complex microbes (38, 39).

The adaptive or neutral models of complexity evolution both fail to capture one of the key aspects of the emergence of biological complexity, namely the striking temporal nonuniformity of the appearance of complex features. The apparent abrupt surges of complexity during evolution are embodied in the concept of major and minor evolutionary transitions (40, 41) and, under a different perspective, in the concept of punctuated equilibrium (42, 43). From a physical standpoint, a rise in complexity is linked to the widespread phenomenon of self-organized criticality (SOC). SOC is a property of dynamical systems with extended degrees of freedom and pronounced nonlinearity whereby the system goes through serial "avalanches" separated in time by intervals of stability. Under the original SOC concept

(44–50), self-similar (power-law) scaling of avalanche sizes is the distinctive feature of the critical dynamics. Here, however, we adopt a more general view of SOC whereby the largest avalanches that appear in the characteristic lifetime of the system are on the scale commensurate with the system size (50).

The behavior of systems with SOC dynamics is obviously reminiscent of the punctuated equilibrium model in evolutionary biology, which involves pronounced temporal discontinuities in evolutionary innovation. This analogy has not escaped the attention of the architects of the SOC concept, who constructed models directly inspired by biological systems and intended to mimic their behavior (46, 47, 49, 50). Specifically, the Bak–Sneppen model (40) employs ecological links between the organisms (represented as the physical proximity in the model space) to drive coevolution of the entire community. Extinction of the least-fit organisms disrupts their local environments and causes concomitant extinction of their closest neighbors. After a short burn-in, such systems self-organize into a critical quasi-equilibrium which is interrupted by avalanches of extinction, with the avalanche size following a power-law distribution.

So far, the exact conditions leading to SOC have not been identified despite considerable effort. However, a notable connection has been shown to exist between competing interactions and frustration in spin glasses, on the one hand, and SOC, on the other hand (51, 52). Specifically, it has been shown that SOC is an emergent property of spin glasses with a diverging number of neighbors (51).

Here, we explore the biological implications of these concepts and findings and conclude that complexity in biological systems emerges in strongly connected multicomponent ensembles, from competing interactions that lead to recurring frustrated states and SOC. We trace these phenomena at all levels of biological organization including both major and minor evolutionary transitions. We further submit that frustration is characteristic of an extremely broad class of hierarchical systems, within and outside biology, and accordingly appears to be the general source of complexity.

Glasses, Patterns, Frustrated States, and SOC

Geometric frustrations in crystallography (10, 53) can be considered the prototype of frustrations in general. Some types of chemical bonds (e.g., Van der Waals and metallic) tend to the closest possible packing of atoms or ions. The local closest packing of hard discs of equal radii in a 2D Euclidean space is provided by the equilateral triangle, and the global closest packing is the triangular lattice built from such triangles. In this case, there is a unique optimal crystal lattice, with no frustrations involved. In the 3D Euclidean space, the optimal local packing is provided by the regular tetrahedron. However, it is impossible to fill the space with the tetrahedra without voids. As a result, there are infinitely many structures corresponding to the same optimal global packing which is less dense than the optimal local packing; such degeneracy is observed also for higher space dimensionalities as well (54).

Another important source of frustrations is competing interactions caused by the coexistence of several types of chemical bonds (in particular, van der Waals and hydrogen, or metallic, covalent, and ionic). Typically, there is no unique optimal structure with the lowest energy; rather, frustrations lead to quasi-degeneracy. Even elemental solids are usually polymorphic, with different phases, such as graphite, diamond, and fullerenes in the case of carbon, between which the energy differences are several orders of magnitude smaller than the energy of each phase. All of the richness and diversity of the structures of minerals and inorganic solids and, as discussed below, of organic molecules and biopolymers come from these frustrations. In particular, the differences in free energies of DNA molecules with different nucleotide sequences are orders of magnitude smaller than the total energy of covalent bonds in these molecules. This flatness of the free energy landscape of DNA provides for the existence of a

combinatorially large number of quasi-degenerate states, that is, genomes of different organisms [Schroedinger's famous aperiodic crystal (55) can be naturally interpreted in these terms].

A notable example is provided by the van der Waals heterostructures (i.e., artificial structures made from different 2D materials); graphene on hexagonal boron nitride (hBN) is the simplest and the best-studied case (56). Graphene and hBN have the same crystal structure but with slightly different lattice periods (1.8% larger for hBN). To minimize the energy of interlayer van der Waals interactions, expansion of graphene to equalize the lattice constants is favorable but such expansion costs some energy of interactions between carbon atoms in graphene. As a result, a distinct pattern is formed at small-enough misalignment angles (57, 58). Similar physics arises and leads to pattern formation when one graphene layer is rotated with respect to another (59). In these cases, patterns originate from the incommensurability of interactions.

Certain types of competing interactions result in the formation of glassy states as captured in the concept of self-induced glassiness (60–62). For example, for magnetic stripe domains, quasi-chaotic patterns result from competition between short-range but strong exchange interaction which tend to maximize magnetization, both locally and globally, and long-range but weak dipole–dipole interaction requiring that the total magnetization of the system be equal to zero (61).

The appearance of the glassy state in extremely simple physical systems with only two competing interactions seems to open the way for understanding the origin of long-term memory, that is, nonergodic processes in which history matters, from the vantage point of statistical mechanics. As first clearly introduced for spin glass by Edwards and Anderson (6), modern physics considers glass to be a distinct state of matter that is intermediate between equilibrium and nonequilibrium (5, 63–68). A characteristic property of glasses is aging, or structural relaxation. Suppose we measure a specific property of an equilibrium phase, liquid or solid, for example the resistivity of a metal (or liquid metal). "Equilibrium" means that, when the measurement is repeated after a thermal cycle (slow heating and cooling down to the initial temperature), we obtain the same value of the resistivity. In glass, the measured value would slowly change from measurement to measurement. The potential energy relief (or landscape, to use a term with biological connotations) for glass is a function with many (asymptotically, infinitely many) local minima separated by barriers with an extremely broad energy distribution. Each local minimum represents a metastable state. During its thermal evolution, the system slowly moves from one minimum to another. Importantly, the glass state is nonergodic: There are many configurations which remain localized in a restricted domain of phase space (63–66). A more-or-less complete formal theory [based on the ideas of Parisi (67)] exists only for an artificial, mean-field-style model with infinitely long-range, independent, random interactions (Sherrington–Kirkpatrick model).

Within this theory, the state of the glass is characterized by an "order parameter" with continuously many components, labeled by a real number $x \in (0,1)$ (67). This number can be represented as an infinite, nonperiodic binary fraction, such as 0.10001110..., where 0 (1) corresponds to the choice of bifurcation on the complex energy relief when cooling down from the equilibrium liquid state. This process of thermal evolution is naturally described in terms of ultrametricity. This is a coarse-grain, "topological" description of the evolution of the system via bifurcations that does not require detailed knowledge of the heights of barriers, rates of transitions, and other characteristics (65). This picture is consistent with the principle of SOC: The SOC dynamics includes avalanches of all sizes distributed by a power law and contains a slow component corresponding to the $1/f$ noise (44). The existence of long-range interactions competing with short-

range ones is essential for the emergence of SOC; in systems with short-range interactions only, SOC is not observed (51).

One of the formal criteria of the glass state is "universal flexibility" (68). Because this is the criterion used in the theory of self-induced glassiness (60–62) that is directly relevant for the present analysis, it merits a brief description. Consider a configuration (of spins, atomic positions, dipolar moments, or other parameters) that is characterized by a function $\varphi(x)$, where x is a d-dimensional vector characterizing a position in space (in most physical applications, $d = 2$ or 3). The energy of this configuration is given by its Hamiltonian $H[\varphi(x)]$ and free energy:

$$F = -T \ln \int D\varphi \exp(-H[\varphi(x)]/T),\qquad [1]$$

where T is the absolute temperature (we put Boltzmann constant equal to one) and $\int D\varphi$ represents summation over all possible configurations. Let us add interaction with another configuration $\sigma(x)$:

$$H[\varphi(x)] \to H_g[\varphi(x)] = H[\varphi(x)] + \frac{g}{2}\int dx[\varphi(x) - \sigma(x)]^2 \qquad [2]$$

and calculate the free energy F_g replacing $H[\varphi(x)] \to H_g[\varphi(x)]$ in Eq. 1. Then, let us consider two transitions: thermodynamic limit $V \to \infty$, where V is the volume of the system, and the limit of infinitely weak coupling $g \to +0$. If these limits do not commute, that is,

$$\lim_{g \to +0}\lim_{V \to \infty}\frac{F_g}{V} \neq \lim_{V \to \infty}\lim_{g \to +0}\frac{F_g}{V} \qquad [3]$$

for a macroscopically large number of configurations $\sigma(x)$, then, the system is glass (i.e., a nonergodic state with memory). Physically, this means that the energy landscape for the glass represents a "universal mapping function," so that for many $\sigma(x)$ there exists a part of the landscape that is minimized by the choice $\varphi(x) = \sigma(x)$. The criterion [3] is fulfilled, under certain conditions, in some nonrandom systems with competing long-range and short-range interactions, such as ferromagnetic thin films (60–62). This criterion appears to be consistent with SOC because, in such systems, the effective connectivity is infinitely large (due to the long-range character of dipole–dipole interactions).

The critical dimensionality, above which the Parisi description holds, is 6 (69). In biological evolutionary dynamics, the dimensionality of the configuration space (fitness landscape) is typically very high; for example, numerous genes in a genome and numerous sites in each gene can be considered separate dimensions (70). Therefore, biological evolution can be expected to follow the chain of causation: competing interactions → frustration → glass-like state → nonergodicity → SOC → evolutionary transitions/"punctuated equilibrium" (Fig. 1). It is important to emphasize that SOC per se is not equivalent to complexity because self-similar structures and phenomena are essentially simple, being conducive to a single mathematical description at all levels. In biology, different levels of hierarchical organization show distinct features such that a universal description is unattainable (i.e., biological systems are genuinely complex and not precisely self-similar). However, as shown in a recent study of holographic complexity in conformal field theories, self-similarity creates optimal conditions for the emergence of complexity, with the maximum rate of complexification caused by a local quench (perturbation of the system, in more general terms) being observed for self-similar systems (71).

When $\sigma(x)$ is simply one specific function, Eq. 3 is equivalent to the condition of spontaneously broken symmetry in the Landau theory of second-order phase transitions (72, 73), where $h(x) = g\sigma(x)$ plays the role of external field conjugated to the
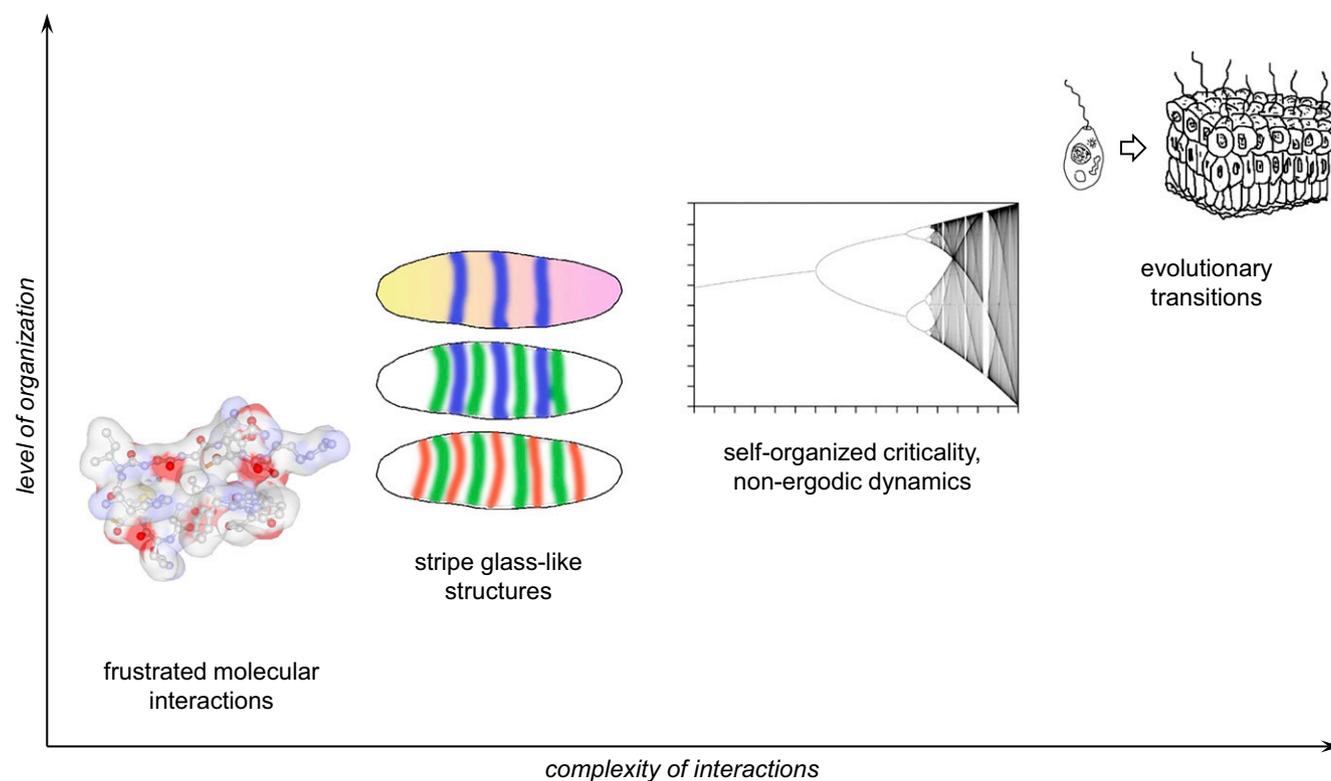
Fig. 1. Competing interactions, frustrated states, SOC, and evolution of complexity.

The figure axes are labeled "level of organization" (vertical) and "complexity of interactions" (horizontal). Labels within the figure read: "frustrated molecular interactions", "stripe glass-like structures", "self-organized criticality, non-ergodic dynamics", and "evolutionary transitions".

order parameter $\varphi(x)$. Conceivably, for some systems, the criteria [**2**] and [**3**] can be satisfied neither for an "almost arbitrary" function $\sigma(x)$ as in glasses nor for a single function as in conventional second-order phase transitions, but rather for a sufficiently rich but limited set of functions. Such systems would spontaneously "glue" to selected configurations from some "library" to form a complex but not completely disordered pattern. Recently, such behavior has been demonstrated in a very simple system, namely, the Ising model with oscillating long-range interactions ("multiwell state") (74). Such patterned glass-like (multiwall) systems are likely to yield better models for biological phenomena than classical glass. Here, we assume that there are many distinct "attractors" and that the energy landscape of the system consists of glass-like parts separated by gaps (74). This model immediately invokes an analogy with pattern recognition in learning theory that has been successfully studied within the framework of the spin-glass formalism (7, 66). A clear and perhaps fundamental analogy from evolutionary biology is a typical, rugged fitness landscape with elevated areas (peaks and plateaus) of high fitness, where an evolving population can travel either upward, under the pressure of selection, or horizontally in a (quasi)neutral evolutionary regime, separated by valleys of low fitness that can be crossed only by genetic drift (thermal fluctuation) (70). In particular, the classical NK model of evolution on rugged fitness landscapes has been developed within the spin-glass framework (75–77).

"Adaptation is localization in sequence space" (78), that is, fixation of distinct genotypes that make evolution proceed along unique trajectories. The feasibility of such fixation depends on the replication fidelity which has to exceed a critical threshold [often called "Eigen threshold," after Manfred Eigen, the originator of the replicator theory (79)] and on the complexity of the fitness landscape. Below the threshold, in a simple, single-peak fitness landscape, a replicator system devolves into a random population of sequences. In physical terms, such a population

represents an ergodic system, whereas an evolving system with fixation on a rugged fitness landscape is nonergodic (Fig. 2). There is a rich-enough set of attractors $\{\sigma(x)\}$ in the sequence space but the standard glass model, where attractors represent a substantial fraction of the configurations so that the landscape is a virtual continuum, does not appear directly relevant for biological evolution. The evolutionary process is made possible by the existence of distinguishable, discrete states. Indeed, genetic information can be changed only in discrete steps, one nucleotide or one codon at a time, at the finest granularity, as opposed to the effective continuity (glass-like character) of the phenotype. This fundamental distinction between the genotype and the phenotype underlies the "central dogma" of molecular biology: The fundamental carriers of biological information (nucleic acids) have to be digital, whereas operational parts, such as proteins, are analog devices (80). Discreteness of biological systems requires that any relevant attractor has a finite basin such that the basins of different attractors should be separated by sufficiently high barriers. RNA molecules can be viewed as occupying an intermediate position between the discreteness of the DNA state and "glassiness" of proteins, being capable of functioning both as information carriers and as operational devices, hence the primordial RNA world scenario. Thus, direct analogies between biology and physics of glasses (11, 12) seem to be oversimplified. The essential concept for biological evolution appears to be that of pattern formation as observed in stripe glasses.

Competing Interactions and Frustrated States as Drivers of Biological Evolution

Competing interactions are apparent among all kinds of biological entities, and frustrated states seem to emerge at all biologically relevant levels of organization, from macromolecules to ecosystems (Table 1). Arguably, the lowest level of specific biological complexity is folding of nucleic acid and protein
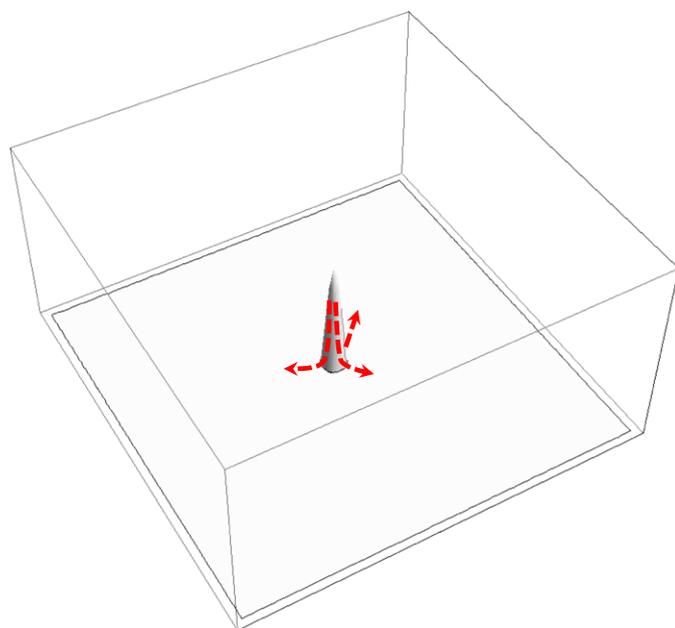
Eigen's single-peak fitness landscape with predictable evolution (survival of the master copy below the error catastrophe threshold or extinction above the threshold)
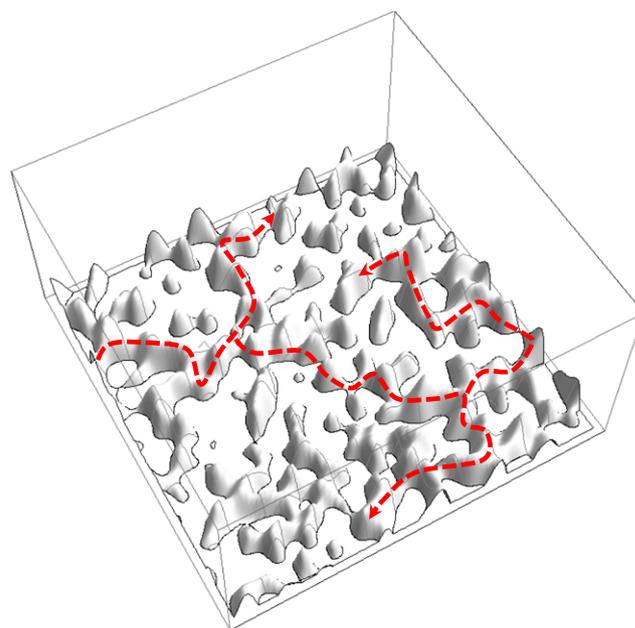
Complex critical percolation fitness landscape with non-ergodic (divergent) evolution

Fig. 2. Complexity of fitness landscapes and nonergodicity of evolution.

molecules into unique, biologically functional 3D structures (81). From the evolutionary perspective, the beginning of life can be most plausibly associated with the appearance of the first RNA molecules (ribozymes) endowed with RNA polymerase activity within the hypothetical, primordial RNA world (82). Not unexpectedly, laboratory experiments that attempt to select for RNA molecules with polymerase activity show that it can be achieved (so far, to a limited extent) only by structurally highly elaborate RNAs (83–85). The competition between short-range and long-range interactions is plainly apparent in protein and RNA molecules and is the defining factor of folding that underlies all molecular functions (14, 15). The transition from the primordial RNA world to the modern-type DNA–RNA–protein biology appears analogous to symmetry breaking whereby the "glassy" landscape of catalytic RNAs is partitioned into the patterned, digital genotype (DNA) and the continuous, analog phenotype (proteins and structural RNAs).

Moving up a level, in macromolecular complexes that, in actuality, perform most, if not all, biochemical functions in cells and viruses (86, 87), the competition between interactions within individual macromolecules and those between subunits that lead to complex formation is equally obvious. Examples abound; suffice it to point out the conformational changes in both ribosomal RNA and proteins during ribosome subunit formation (88, 89), in transcription factors upon DNA binding (90, 91), and in virion proteins during morphogenesis of virus particles (92). Furthermore, allosteric regulation of enzymatic complexes involves transition between macromolecular conformations with close free energies but distinct biological properties (81). The existence of many conformations with comparable energies in biological complexes has the same cause as the polymorphism of simple (including elemental) inorganic solids mentioned above, namely, competing interactions.

On another plane of biological organization that is unique to living matter, competing interactions can be conceptualized as selection pressures that act in opposite directions and produce various tradeoffs that permeate biology (93, 94). A complex fitness landscape with many peaks and basins of attraction (that is, with many evolutionary strategies with more or less the same adaptive efficiency) results from an interplay of such competing factors. In particular, such conflicting selective processes are a key, inherent component of host–parasite coevolution. Emergence of genetic parasites appears to be inevitable in all evolving replicator systems because it can be shown that parasite-protected systems are evolutionarily unstable (95, 96). Moreover, genetic parasites with different reproduction strategies (viruses, plasmids, transposons, and more) accompany (nearly) all cellular life forms (97–100). The frustrated state of a host–parasite system is caused by a complex interplay between the parasite replication, the host replication, and the interactions that stabilize the host–parasite system as a whole. The conflicts between the selective factors that operate on each of these levels appear to be a major, perhaps the principal, driver of evolution of these systems (101). Computer simulations under a wide range of conditions show that, in a well-mixed replicator system, parasites overwhelm the hosts, resulting in the eventual collapse of the entire host–parasite system (96, 102–105). In contrast, compartmentalization (that, in specific terms, could represent, for instance, partitioning of replicator ensembles between different microcompartments in an inorganic compartment network or simply separation in a viscous medium) stabilizes the system and leads to diversification and evolution of complexity (102–104). Effectively, the outcome of host–parasite coevolution in such modeling studies is pattern formation, a typical consequence of frustration in glass-like states.

Compartmentalization is arguably the simplest, most fundamental effect of host–parasite conflicts but, in all cellular life forms, these conflicts also drive the evolution of versatile host

Table 1. Competing interactions and frustrated states in biological evolution

System	Frustration-producing factors (competing interactions)	Emergent functional and evolutionary features
RNA	Short-range (within stem local hydrogen bonding, stacking) vs. long-range (long-distance hydrogen bonding, salt bridges) interactions between nucleotides	Complex 3D structures including ribozymes
Proteins	Short-range (Van der Waals) vs. long-range (hydrogen bonds, salt bridges) interactions between amino acid side chains	Stable conformations and semiregular patterns in protein structures; allostery enabled by transitions between energetically quasi-degenerate conformations
Macromolecular complexes	Within-subunit vs. between-subunit interactions	Elaborate complex organization, in particular nucleoproteins (ribosomes, chromatin)
Cells	**Membranes (confinement of chemicals) vs. channels/pores (transport of chemicals)**	**Compartments and cellular machinery dependent on electrochemical gradients**
Autonomous (hosts) and semiautonomous (parasites) replicators	**Replicator vs. parasite genomes**	**Self- vs. non-self-discrimination and defense; complex genomes of increasing size; primitive cells**
Autonomous (hosts) and semiautonomous (parasites) reproducers/replicators	**Host cells and viruses**	**Infection mechanisms, defense and counterdefense systems, evolutionary arms race; contribution to the origin of multicellular life forms**
Autonomous (hosts) and semiautonomous (parasites) reproducers/replicators	Host cells vs. transposons	Intragenomic DNA replication control; evolutionary innovation through recruitment of transposon sequences
Autonomous (hosts) and semiautonomous (parasites) reproducers/replicators	Host cells vs. plasmids	Beneficial cargo genes, plasmid addiction systems, efficient gene exchange and transfer mechanisms
Emerging eukaryotic cells	**Host (archaeal) cells vs. endosymbiont (α-proteobacteria, protomitochondria)**	**Eukaryotic cells with complex intracellular organization**
Communities of unicellular organisms	**Individual cells vs. cellular ensembles**	**Information exchange and quorum sensing mechanisms; replication control, programmed cell death, multicellularity**
Multicellular organisms	Soma vs. germline	Complex bodies, tissues and organ differentiation, sexual reproduction
Multicellular organisms	Dividing vs. quiescent cells	Aging, cancer, death
Populations	**Individual members vs. groups**	**Population-level cooperation; kin selection; eusocilaity**
Populations	Males vs. females (partners with unequal parental investment)	Sexual selection, sexual dimorphism
Ecosystems	Species in different niches	Interspecies competition, host–parasite and predator–prey relationships, mutualism, symbiosis
Societies*	**—**	**—**

Those competing interactions and frustrated states that are deemed to directly contribute to MTE are shown in bold.
*We refrain from specifying the conflicts that drive the origin and evolution of human societies.

defense systems and counterdefense systems in parasites, another prominent and ubiquitous manifestation of biological complexity (106–109). In the course of evolution, conflicts between hosts and parasites are resolved into multiple, distinct, stable evolutionary regimes. These regimes span the entire range of host–parasite relationships, from highly aggressive parasites, such as lytic viruses that kill the host and move to the next one, to cooperative elements, such as many plasmids that often provide beneficial functionalities to the host (99, 100). This persistent diversity of host–parasite interactions is a major part of biological complexity at the level of ecosystems and the entire biosphere.

Even more generally, it seems to be no exaggeration to state that frustration caused by intergenomic conflicts drives the evolution of all biological complexity (101, 110). The genomes of all cellular life forms contain multiple inserted MGE genomes which in multicellular eukaryotes (animals and plants) account for the majority of the genome sequence (111–113). The competing interactions causing frustration are particularly obvious in the case of MGE with dual activities, such as toxin–antitoxin (TA) and restriction-modification (RM) modules in prokaryotes (114–117). On the one hand, TA and RM systems protect the hosts from other, more aggressive parasites, primarily viruses, but on the other hand they themselves behave as MGE. Frustration in bacteria and archaea that harbor TA and RM (that is, nearly all bacteria and archaea) is manifest at several levels. The TA and RM systems compete, on the one hand, with viruses, which they attack and hence protect the host, and on the other hand with the host itself, which they kill when it "attempts" to get rid of these elements. Another part of this gamut are plasmids on which RM and TA modules are typically transferred. The outcome of these complex networks of competition is the stabilization of the entire host–parasite system in which components with all types of reproduction strategies persist indefinitely. In other words, host–parasite coexistence translates into persistence of biological complexity at the ecosystem level. A striking feature of the competition networks is the "guns for hire" phenomenon, that is, shuttling of the same active components, such as nucleases involved in transposition, between defense systems and MGE (118, 119).

EVOLUTION

Evolutionary Transitions and Major Innovations Driven by Competing Interactions

Competing interactions and frustration naturally apply to evolutionary transitions and, more generally, major evolutionary innovations. The concept of major transitions in evolution (MTEs) developed by Maynard Smith and Szathmáry (40, 41) defines a distinct class of major evolutionary innovations that involve evolutionary transitions in individuality. A classic example of an MTE is the origin of multicellular organisms from unicellular life forms, but MTEs, although not numerous, punctuate the entire history of life (Table 1). The key tenet of the MTE concept is that, within its framework, the transitions are construed not simply as major innovations but rather meet strict criteria that make them akin to phase transitions in physics. Thus, every MTE is a major innovation, but not vice versa. The signature feature of MTEs is an evolutionary transition in individuality which involves a change in the level of selection, for example from a single cell to an ensemble of cell (a multicellular organism) (Fig. 1 and Table 1). The second signature of MTEs, *sensu* Szathmáry (41), is that each transition is associated with the emergence of a new type of information storage, use, and transmission (e.g., multicellularity is linked to the rise of epigenetic informational systems). Competing interactions and/or levels of selections are immediately apparent for each MTE (Table 1).

Starting from the most obvious, evolution of multicellularity involves the intrinsic conflict between selection forces acting at the level of individual cells and those that are manifest at the level of cellular ensembles or tissues. Clearly, to maintain the functionality of a multicellular organism, the proliferation of individual cells has to be tightly controlled.

Moving back in time, we know little about the origin of the first cells. It is nevertheless difficult to imagine an evolutionary scenario in which the emergence of cells was not preceded by an evolutionary stage at which all genetic information was encoded in small elements that would resemble modern MGE (120–122). Subsequent evolution involved accretion of such elements to form large genomes such as those of modern prokaryotes. Under this scenario, the emergence of cells involved competition between the selection factors that affect individual genetic elements and those that act on ensembles of such elements that formed cellular genomes. Once again, replication of individual elements has to be subdued for the ensemble (i.e., the cell) to function. A closely similar evolutionary scenario has been implemented in a recent mathematical model of primordial cell evolution (123).

The next MTE, the origin of eukaryotes, is associated with endosymbiosis that gave rise to the mitochondria and hence involved the inevitable conflict between the endosymbiont and the evolving eukaryotic cell that required coordinated reproduction of the host and the symbiont to survive and function (124–126). The same conflict is inherent in the evolution of photosynthesizing algae via the cyanobacterial endosymbiosis that gave rise to the chloroplast. The frustration caused by host–symbiont conflicts in these MTEs was resolved by the formation of the stable symbiotic associations, but the conflicts linger, for example in the form of mitochondrial diseases (127) and frequent lysis of mitochondria which in some organisms results in frequent insertion of nonfunctional mitochondrial DNA into the host genome (128).

The later MTEs that led to the emergence of eusocial animals and societies also clearly involve competition between individuals and collective, or between collectives at different levels of organization. Generally, it appears that for evolutionary transitions in individuality, which define MTEs, competing interactions between entities at different organizational levels are the intrinsic driving factor.

Notably, the conflicts between the different levels of individuality in MTEs appear to be intertwined with host–parasite conflicts. Indeed, mathematical models of the evolution of multicellularity suggest that defense against viruses could be a major driving factor of this MTE. In particular, parasite pressure drives the evolution of programmed cell death, a distinct form of defense that functions only in the presence of cell aggregation and appears to promote emergence of multicellularity (129). Genetic parasites appear to have played important roles also in earlier MTEs, in particular the origin of eukaryotes that is thought to have been accompanied by massive invasion of introns from the endosymbiont into the host genome (125, 126, 130). This explosive invasion of parasitic elements into the genomes of the emerging eukaryotes could have triggered the evolution of several key features of eukaryotic cells that are central to the dramatic increase in the complexity of the cellular organization compared with the prokaryotic ancestors, including the nucleus, the spliceosome, and the ubiquitin signaling network (131). Although the exact sequence of events in eukaryogenesis is difficult to ascertain, defense against the deleterious effect of the invading introns, in particular production of aberrant proteins from unspliced transcripts, is likely to be a major selective force behind the evolution of these signatures of the eukaryotic cellular organization (125, 126, 131). However meager our knowledge of the origins of the first cells might be, it appears all but certain that conflicts between selfish and "cooperative" genetic elements played an important role (120, 122). Thus, all MTEs seem to involve more than one type of competing interactions.

Every MTE involves major evolutionary innovations, but not every major innovation is an MTE. Nevertheless, it can be argued that MTEs differ from other evolutionary innovations in degree rather than in kind. The major evolutionary innovations that are not associated with a change in the level of selection nevertheless involve local transitions in individuality, that is, emergence of new, complex functions through evolutionary fixation of new interactions between genes, such as photosystems components in the case of photosynthesis (132–134) or enzymes of methanogenic pathways in the case of archaeal methanogenesis (135, 136). It can be argued that all evolutionary innovations involve emergence of new units of selection (Darwinian individual), if not a new level of selection (class of Darwinian individuals). Accordingly, it seems that competing interactions and frustration are inherent in all major evolutionary innovations.

It seems important to note that the MTEs, and more generally evolutionary innovations, can also be naturally interpreted in terms of cooperation between Darwinian individuals that leads to the emergence of a higher level of selection. Crucially, competing interactions directly lead to cooperation, as most clearly illustrated by the evolution of multicellularity as a result of parasite pressure and trade-offs between hoarding and public goods production (101, 129). Succinctly put, competition begets cooperation.

Cancer, Aging, and Death

The conflict between the propagation of individual cells and the maintenance of stable tissues and organs in multicellular organisms is resolved via multilayer systems of controls of cell proliferation, another striking manifestation of biological complexity. However, an alternative and common resolution of this conflict triggered by impairment of such mechanisms is unchecked proliferation of cheater cells that can lead to tumorigenesis and, in particular, cancer in animals (137). The emergence of cheaters in cell ensembles appears inescapable for the same reasons that make the emergence of parasites intrinsic to any replicator system (95).

Furthermore, aging, apparently an inherent feature of multicellular organisms, seems to be caused by the same conflict (138). This is the case because elimination of senescent cells that have accumulated deleterious somatic mutations via competition with high-fitness cells conflicts with the propensity of the latter to divide uncontrollably and thus form tumors. In other words, there is an inherent conflict between the fitness ("vigor") of an individual cell that consists in high division rate and cellular cooperation that requires limiting that rate or eliminating division altogether.

Frustration as the Key Factor Underlying All Complexity in Nature and the Specifics of Biological Evolution

Although unifying explanations of universal phenomena are inherently dangerous, it does appear plausible that complexity can emerge only in nonergodic systems, and nonergodicity is caused by competing interactions. Furthermore, the competition between short-range and long-range interactions can, under additional constraints, result in SOC, which provides the conditions for the evolution of complexity. This general perspective on the origin of complexity seems to explain all types of patterns existing in nature, from stripe glasses to planetary systems and galaxies (Fig. 1). However, contrary to some general statements (50), there is more to the evolution of complexity than SOC. Importantly, SOC does not lead to hierarchical complexity: Fractal patterns produced by SOC are not genuinely complex because, by definition, they appear the same at all spatial and/or temporal scales. It should be noted that, as an alternative to SOC, evolution of biological complexity has been explored in abstract models based on the concept of highly optimized tolerance (HOT) which emerges through robustness trade-offs, in particular competition between generalists and specialists in uncertain environments (139, 140). A detailed analysis and discussion of HOT is beyond the scope of this article, but it appears that frustrations caused by competing evolutionary factors, such as adaptation to either a broad or a narrow range of conditions, underlie this model as well.

In biology, the patterns observed on different levels of organization, such as macromolecules, cells, multicellular organisms, populations, and ecosystems, are different. These levels of organization are linked through MTEs which, as argued above, are driven by competing interactions and the resulting frustration. Indeed, the main message of the present work is that this appears to be the universal mechanism driving the evolution of hierarchy in nature. Here, we do not present the actual mathematical theory of frustration-driven evolution. It seems likely that physical and mathematical ideas and formalisms beyond those currently known are necessary to develop such a theory (141).

What are the specific, defining features of life? The distinction of biological systems is certainly not that "nothing in biology makes sense except in the light of evolution" (142), whereas in the rest of the world "everything does." As discussed above, outside of the quantum realm, the world is full of nonergodic systems, and conceivably the universe itself can be properly described only in the light of its evolution over the 13.8 billion y since the Big Bang (143). However, the level of complexity and elaboration characteristic of living things is unmatched by anything outside biology. The key difference between biological entities and inanimate nonergodic, complex systems seems to be the unique biological memory mechanism combined with the phenotype-to-genotype mapping. Specifically, this mechanism involves replication of digital information carriers (nucleic acids) that memorize the patterns emerging from competing interactions at different levels and propagate that memory with sufficient fidelity to allow selection via the phenotype-to-genotype feedback. Attempts to define life at a fundamental level could be viewed as philosophical exercises of limited interest (144–146). Nevertheless, it does appear that complexity emerging from competing interactions, combined with memory perpetuated by replication of information carriers, underlies all life, and, conversely, any system endowed with these properties will qualify as living.

1. Feynman RP (2017) *The Character of Physical Law* (MIT Press, Cambridge, MA).
2. Smith TF, Morowitz HJ (1982) Between history and physics. *J Mol Evol* 18:265–282.
3. Goldenfeld N, Woese CR (2011) Life is physics: Evolution as a collective phenomenon far from equilibrium. *Annu Rev Condens Matter Phys* 2:375–399.
4. Newton I (1999) *The Principia. The Mathematical Principles of Natural Philosophy* (Univ of California Press, Berkeley, CA).
5. Einstein A, Infeld L (1967) *The Evolution of Physics* (Simon and Schuster, New York).
6. Edwards SF, Anderson PW (1975) Theory of spin glasses. *J Phys F Met Phys* 5:965–974.
7. Fischer KH, Hertz JA (1993) *Spin Glasses* (Cambridge Univ Press, Cambridge, UK).
8. Kubo R, Toda M, Hashitsume N (1985) *Statistical Physics II: Nonequilibrium Statistical Mechanics* (Springer, New York).
9. Jin F, et al. (2013) Equilibration and thermalization of classical systems. *New J Phys* 15:033009.
10. Sadoc JF, Mozzeri R (1999) *Geometric Frustration* (Cambridge Univ Press, Cambridge, UK).
11. Laughlin RB, Pines D (2000) The theory of everything. *Proc Natl Acad Sci USA* 97:28–31.
12. Laughlin RB, Pines D, Schmalian J, Stojkovic BP, Wolynes P (2000) The middle way. *Proc Natl Acad Sci USA* 97:32–37.
13. Manyuhina OV, Fasolino A, Katsnelson MI (2007) Slow dynamics in a model of the cellulose network. *Polymer (Guelph)* 48:4649–4854.
14. Ferreiro DU, Komives EA, Wolynes PG (2014) Frustration in biomolecules. *Q Rev Biophys* 47:285–363.
15. Ferreiro DU, Komives EA, Wolynes PG (2018) Frustration, function and folding. *Curr Opin Struct Biol* 48:68–73.
16. Adami C (2002) What is complexity? *BioEssays* 24:1085–1094.
17. Adami C (2012) The use of information theory in evolutionary biology. *Ann N Y Acad Sci* 1256:49–65.
18. Adami C, Ofria C, Collier TC (2000) Evolution of biological complexity. *Proc Natl Acad Sci USA* 97:4463–4468.
19. Gell-Mann M (1995) *The Quark and the Jaguar: Adventures in the Simple and the Complex* (St. Martin's Griffin, New York).
20. McShea DW, Brandon RN (2010) *Biology's First Law: The Tendency for Diversity and Complexity to Increase in Evolutionary Systems* (Univ of Chicago Press, Chicago).
21. Koonin EV (2016) The meaning of biological information. *Philos Trans R Soc A* 374:20150065.
22. Koonin EV (2004) A non-adaptationist perspective on evolution of genomic complexity or the continued dethroning of man. *Cell Cycle* 3:280–285.
23. Marshall SM, Murray ARG, Cronin L (2017) A probabilistic framework for identifying biosignatures using pathway complexity. *Philos Trans R Soc A* 375:20160342.
24. Gell-Mann M, Lloyd S (1996) Information measures, effective complexity, and total information. *Complexity* 2:44–52.
25. Badii R, Politi A (1997) *Complexity. Hierarchical Structures and Scaling in Physics* (Cambridge Univ Press, Cambridge, UK).
26. Lloyd S (2001) Measures of complexity: A nonexhaustive list. *IEEE Contr Syst* 21:7–8.
27. Heim NA, et al. (2017) Hierarchical complexity and the size limits of life. *Proc R Soc B* 284:20171039.
28. Gould SJ, Lewontin RC (1979) The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proc R Soc Lond B Biol Sci* 205:581–598.
29. Gould SJ (1997) The exaptive excellence of spandrels as a term and prototype. *Proc Natl Acad Sci USA* 94:10750–10755.
30. Lynch M (2006) The origins of eukaryotic gene structure. *Mol Biol Evol* 23:450–468.
31. Lynch M (2007) The frailty of adaptive hypotheses for the origins of organismal complexity. *Proc Natl Acad Sci USA* 104:8597–8604.
32. Lynch M (2007) *The Origins of Genome Archicture* (Sinauer, Sunderland, MA).
33. Lynch M, Conery JS (2003) The origins of genome complexity. *Science* 302:1401–1404.
34. Stoltzfus A (1999) On the possibility of constructive neutral evolution. *J Mol Evol* 49:169–181.
35. Stoltzfus A (2012) Constructive neutral evolution: Exploring evolutionary theory's curious disconnect. *Biol Direct* 7:35.
36. Gray MW, Lukes J, Archibald JM, Keeling PJ, Doolittle WF (2010) Cell biology. Irremediable complexity? *Science* 330:920–921.
37. Lukeš J, Archibald JM, Keeling PJ, Doolittle WF, Gray MW (2011) How a neutral evolutionary ratchet can build cellular complexity. *IUBMB Life* 63:528–537.
38. Sela I, Wolf YI, Koonin EV (2016) Theory of prokaryotic genome evolution. *Proc Natl Acad Sci USA* 113:11399–11407.

EVOLUTION

39. Iranzo J, Cuesta JA, Manrubia S, Katsnelson MI, Koonin EV (2017) Disentangling the effects of selection and loss bias on gene dynamics. *Proc Natl Acad Sci USA* 114: E5616–E5624.

40. Maynard Smith J, Szathmáry E (1997) *The Major Transitions in Evolution* (Oxford Univ Press, Oxford).

41. Szathmáry E (2015) Toward major evolutionary transitions theory 2.0. *Proc Natl Acad Sci USA* 112:10104–10111.

42. Gould SJ (1994) Tempo and mode in the macroevolutionary reconstruction of Darwinism. *Proc Natl Acad Sci USA* 91:6764–6771.

43. Eldredge N, Gould SJ (1997) On punctuated equilibria. *Science* 276:338–341.

44. Bak P, Tang C, Wiesenfeld K (1987) Self-organized criticality: An explanation of the 1/f noise. *Phys Rev Lett* 59:381–384.

45. Bak P, Tang C, Wiesenfeld K (1988) Self-organized criticality. *Phys Rev A Gen Phys* 38: 364–374.

46. Bak P, Sneppen K (1993) Punctuated equilibrium and criticality in a simple model of evolution. *Phys Rev Lett* 71:4083–4086.

47. Maslov S, Paczuski M, Bak P (1994) Avalanches and 1/f noise in evolution and growth models. *Phys Rev Lett* 73:2162–2165.

48. Maslov S, Zhang YC (1995) Exactly solved model of self-organized criticality. *Phys Rev Lett* 75:1550–1553.

49. Bak P, Paczuski M (1995) Complexity, contingency, and criticality. *Proc Natl Acad Sci USA* 92:6689–6696.

50. Bak P (1996) *How Nature Works: The Science of Self-Organized Criticality* (Springer, New York).

51. Andresen JC, et al. (2013) Self-organized criticality in glassy spin systems requires a diverging number of neighbors. *Phys Rev Lett* 111:097203.

52. Pázmándi F, Zaránd G, Zimányi GT (1999) Self-organized criticality in the hysteresis of the Sherrington-Kirkpatrick model. *Phys Rev Lett* 83:1034–1037.

53. Nelson DR (1983) Order, frustration, and defects in liquids and glasses. *Phys Rev B* 28: 5515–5535.

54. Conway JH, Sloane NJA (1988) *Sphere Packings, Lattices and Groups* (Springer, New York).

55. Schroedinger E (1944) *What is Life? The Physical Aspect of the Living Cell* (Trinity College Press, Dublin).

56. Geim AK, Grigorieva IV (2013) Van der Waals heterostructures. *Nature* 499:419–425.

57. Woods CR, et al. (2014) Commensurate-incommensurate transition in graphene on hexagonal boron nitride. *Nat Phys* 10:451–456.

58. van Wijk MM, Schuring A, Katsnelson MI, Fasolino A (2014) Moiré patterns as a probe of interplanar interactions for graphene on h-BN. *Phys Rev Lett* 113:135504.

59. van Wijk MM, Schuring A, Katsnelson MI, Fasolino A (2015) Relaxation of moire patterns for slightly misaligned identical lattices: Graphene on graphite. *2D Mater* 2: 034010.

60. Schmalian G, Wolynes PG (2000) Stripe glasses: Self-generated randomness in a uniformly frustrated system. *Phys Rev Lett* 85:836–839.

61. Principi A, Katsnelson MI (2016) Stripe glasses in ferromagnetic thin films. *Phys Rev B* 93:054410.

62. Principi A, Katsnelson MI (2016) Self-induced glassiness and pattern formation in spin systems subject to long-range interactions. *Phys Rev Lett* 117:137201.

63. Binder K, Young AP (1986) Spin glasses: Experimental facts, theoretical concepts, and open questions. *Rev Mod Phys* 58:801–976.

64. Das SP (2004) Mode-coupling theory and the glass transition in supercooled liquids. *Rev Mod Phys* 76:785–851.

65. Rammal R, Toulouse G, Virasoro MA (1986) Ultrametricity for physicists. *Rev Mod Phys* 58:765–788.

66. Mezard M, Parisi G, Virasoro MA, eds (1987) *Spin Glass Theory and Beyond* (World Scientific, Singapore).

67. Parisi G (1980) A sequence of approximated solutions to the S-K model for spin glasses. *J Phys A Math Gen* 13:1101–1112.

68. Monasson R (1995) Structural glass transition and the entropy of the metastable states. *Phys Rev Lett* 75:2847–2850.

69. Wang W, Moore MA, Katzgraber HG (2017) Fractal dimension of interfaces in Edwards-Anderson spin glasses for up to six space dimensions. *Phys Rev E* 97:032104.

70. Gavrilets S (2004) *Fitness Landscapes and the Origin of Species* (Princeton Univ Press, Princeton).

71. Ageev DS, Aref'eva IY, Bagrov AA, Katsnelson MI (August 14, 2018) Holographic local quench and effective complexity. *J High Energy Phys*, 10.1007/JHEP08(2018)071.

72. Landau LD, Lifshitz EM (1980) *Statistical Physics* (Pergamon, Oxford).

73. Ruelle D (1999) *Statistical Mechanics: Rigorous Results* (World Scientific, Singapore).

74. Kolmus A, Katsnelson MI, Khajetoorians AA, Kappen HJ (2018) Tunable glassiness on a two-dimensional atomic spin array. arXiv:1804.08093.

75. Kauffman SA, Weinberger ED (1989) The NK model of rugged fitness landscapes and its application to maturation of the immune response. *J Theor Biol* 141:211–245.

76. Kauffman SA, Macready WG (1995) Search strategies for applied molecular evolution. *J Theor Biol* 173:427–440.

77. Stein DL (1992) *Spin Glasses and Biology* (World Scientific, Singapore).

78. Nowak MA (2006) *Evolutionary Dynamics: Exploring the Equations of Life* (Belknap, Cambridge, MA).

79. Eigen M (1971) Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* 58:465–523.

80. Koonin EV (2015) Why the central dogma: On the nature of the great biological exclusion principle. *Biol Direct* 10:52.

81. Finkelstein AV, Ptitsyn OB (2002) *Protein Physics: A Course of Lectures* (Academic, New York).

82. Atkins JF, Gesteland RF, Cech TR, eds (2010) *RNA Worlds: From Life's Origins to Diversity in Gene Regulation* (Cold Spring Harbor Lab Press, Cold Spring Harbor, NY).

83. McGinness KE, Joyce GF (2003) In search of an RNA replicase ribozyme. *Chem Biol* 10: 5–14.

84. Horning DP, Joyce GF (2016) Amplification of RNA by an RNA polymerase ribozyme. *Proc Natl Acad Sci USA* 113:9786–9791.

85. Samanta B, Joyce GF (2017) A reverse transcriptase ribozyme. *eLife* 6:e31153.

86. Gavin AC, et al. (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415:141–147.

87. Krause R, von Mering C, Bork P (2003) A comprehensive set of protein complexes in yeast: Mining large scale protein-protein interaction screens. *Bioinformatics* 19: 1901–1908.

88. Shajani Z, Sykes MT, Williamson JR (2011) Assembly of bacterial ribosomes. *Annu Rev Biochem* 80:501–526.

89. Kim H, et al. (2014) Protein-guided RNA dynamics during early ribosome assembly. *Nature* 506:334–338.

90. Wolberger C (1999) Multiprotein-DNA complexes in transcriptional regulation. *Annu Rev Biophys Biomol Struct* 28:29–56.

91. Ogata K, Sato K, Tahirov TH (2003) Eukaryotic transcriptional regulatory complexes: Cooperativity from near and afar. *Curr Opin Struct Biol* 13:40–48, and erratum (2003) 13:262.

92. Perlmutter JD, Hagan MF (2015) Mechanisms of virus assembly. *Annu Rev Phys Chem* 66:217–239.

93. Poelwijk FJ, de Vos MG, Tans SJ (2011) Tradeoffs and optimality in the evolution of gene regulation. *Cell* 146:462–470.

94. de Vos MG, Dawid A, Sunderlikova V, Tans SJ (2015) Breaking evolutionary constraint with a tradeoff ratchet. *Proc Natl Acad Sci USA* 112:14906–14911.

95. Koonin EV, Wolf YI, Katsnelson MI (2017) Inevitability of the emergence and persistence of genetic parasites caused by thermodynamic instability of parasite-free states. *Biol Direct* 12:31.

96. Szathmáry E (2000) The evolution of replicators. *Philos Trans R Soc Lond B Biol Sci* 355:1669–1676.

97. Holmes EC (2009) *The Evolution and Emergence of RNA Viruses* (Oxford Univ Press, Oxford).

98. Koonin EV (2011) *The Logic of Chance: The Nature and Origin of Biological Evolution* (FT, Upper Saddle River, NJ).

99. Koonin EV, Starokadomskyy P (2016) Are viruses alive? The replicator paradigm sheds decisive light on an old but misguided question. *Stud Hist Philos Biol Biomed Sci* 59:125–134.

100. Jalasvuori M, Koonin EV (2015) Classification of prokaryotic genetic replicators: Between selfishness and altruism. *Ann N Y Acad Sci* 1341:96–105.

101. Koonin EV (2016) Viruses and mobile elements as drivers of evolutionary transitions. *Philos Trans R Soc B* 371:20150442.

102. Takeuchi N, Hogeweg P (2008) Evolution of complexity in RNA-like replicator systems. *Biol Direct* 3:11.

103. Takeuchi N, Hogeweg P (2012) Evolutionary dynamics of RNA-like replicator systems: A bioinformatic approach to the origin of life. *Phys Life Rev* 9:219–263.

104. Takeuchi N, Hogeweg P, Koonin EV (2011) On the origin of DNA genomes: Evolution of the division of labor between template and catalyst in model replicator systems. *PLOS Comput Biol* 7:e1002024.

105. Szathmáry E, Demeter L (1987) Group selection of early replicators and the origin of life. *J Theor Biol* 128:463–486.

106. Labrie SJ, Samson JE, Moineau S (2010) Bacteriophage resistance mechanisms. *Nat Rev Microbiol* 8:317–327.

107. Stern A, Sorek R (2011) The phage-host arms race: Shaping the evolution of microbes. *BioEssays* 33:43–51.

108. Makarova KS, Wolf YI, Koonin EV (2013) Comparative genomics of defense systems in archaea and bacteria. *Nucleic Acids Res* 41:4360–4377.

109. Koonin EV, Makarova KS, Wolf YI (2017) Evolutionary genomics of defense systems in archaea and bacteria. *Annu Rev Microbiol* 71:233–261.

110. Aravind L, Anantharaman V, Zhang D, de Souza RF, Iyer LM (2012) Gene flow and biological conflict systems in the origin and evolution of eukaryotes. *Front Cell Infect Microbiol* 2:89.

111. Kazazian HH, Jr (2004) Mobile elements: Drivers of genome evolution. *Science* 303: 1626–1632.

112. Oliver KR, Greene WK (2009) Transposable elements: Powerful facilitators of evolution. *BioEssays* 31:703–714.

113. Chénais B, Caruso A, Hiard S, Casse N (2012) The impact of transposable elements on eukaryotic genomes: From genome size increase to genetic adaptation to stressful environments. *Gene* 509:7–15.

114. Gerdes K, Christensen SK, Løbner-Olesen A (2005) Prokaryotic toxin-antitoxin stress response loci. *Nat Rev Microbiol* 3:371–382.

115. Makarova KS, Wolf YI, Koonin EV (2009) Comprehensive comparative-genomic analysis of type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. *Biol Direct* 4:19.

116. Van Melderen L (2010) Toxin-antitoxin systems: Why so many, what for? *Curr Opin Microbiol* 13:781–785.

117. Van Melderen L, Saavedra De Bast M (2009) Bacterial toxin-antitoxin systems: More than selfish entities? *PLoS Genet* 5:e1000437.

118. Koonin EV, Krupovic M (January 1, 2015) A movable defense. *The Scientist*. Available at https://www.the-scientist.com/features/a-movable-defense-36135.

119. Koonin EV, Makarova KS (2017) Mobile genetic elements and evolution of CRISPR-Cas systems: All the way there and back. *Genome Biol Evol* 9:2812–2825.

120. Koonin EV, Martin W (2005) On the origin of genomes and cells within inorganic compartments. *Trends Genet* 21:647–654.

121. Koonin EV, Senkevich TG, Dolja VV (2006) The ancient virus world and evolution of cells. *Biol Direct* 1:29.

122. Koonin EV (2009) On the origin of cells and viruses: Primordial virus world scenario. *Ann N Y Acad Sci* 1178:47–64.

123. Takeuchi N, Hogeweg P, Kaneko K (2017) The origin of a primordial genome through spontaneous symmetry breaking. *Nat Commun* 8:250.

124. Embley TM, Martin W (2006) Eukaryotic evolution, changes and challenges. *Nature* 440:623–630.

125. Martin W, Koonin EV (2006) Introns and the origin of nucleus-cytosol compart-mentalization. *Nature* 440:41–45.

126. López-García P, Moreira D (2006) Selective forces for the origin of the eukaryotic nucleus. *BioEssays* 28:525–533.

127. Suomalainen A, Battersby BJ (2018) Mitochondrial diseases: The contribution of organelle stress responses to pathology. *Nat Rev Mol Cell Biol* 19:77–92.

128. Bensasson D, Zhang D, Hartl DL, Hewitt GM (2001) Mitochondrial pseudogenes: Evolution's misplaced witnesses. *Trends Ecol Evol* 16:314–321.

129. Iranzo J, Lobkovsky AE, Wolf YI, Koonin EV (2014) Virus-host arms race at the joint origin of multicellularity and programmed cell death. *Cell Cycle* 13:3083–3088.

130. Koonin EV (2009) Intron-dominated genomes of early ancestors of eukaryotes. *J Hered* 100:618–623.

131. Koonin EV (2006) The origin of introns and their role in eukaryogenesis: A compromise solution to the introns-early versus introns-late debate? *Biol Direct* 1:22.

132. Raymond J, Zhaxybayeva O, Gogarten JP, Gerdes SY, Blankenship RE (2002) Whole-genome analysis of photosynthetic prokaryotes. *Science* 298:1616–1620.

133. Mulkidjanian AY, et al. (2006) The cyanobacterial genome core and the origin of photosynthesis. *Proc Natl Acad Sci USA* 103:13126–13131.

134. Nowicka B, Kruk J (2016) Powered by light: Phototrophy and photosynthesis in prokaryotes and its evolution. *Microbiol Res* 186–187:99–118.

135. Martin WF, Sousa FL (2015) Early microbial evolution: The age of anaerobes. *Cold Spring Harb Perspect Biol* 8:a018127.

136. Sorokin DY, et al. (2017) Discovery of extremely halophilic, methyl-reducing eur-yarchaea provides insights into the evolutionary origin of methanogenesis. *Nat Microbiol* 2:17081.

137. Hanahan D, Weinberg RA (2011) Hallmarks of cancer: The next generation. *Cell* 144:646–674.

138. Nelson P, Masel J (2017) Intercellular competition and the inevitability of multicel-lular aging. *Proc Natl Acad Sci USA* 114:12982–12987.

139. Carlson JM, Doyle J (1999) Highly optimized tolerance: A mechanism for power laws in designed systems. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* 60:1412–1427.

140. Zhou T, Carlson JM, Doyle J (2005) Evolutionary dynamics and highly optimized tolerance. *J Theor Biol* 236:438–447.

141. Katsnelson MI, Wolf YI, Koonin EV (2018) Towards physical principles of biological evolution. *Phys Scr* 93:043001.

142. Dobzhansky T (1973) Nothing in biology makes sense except in the light of evolu-tion. *Am Biol Teach* 35:125–129.

143. Weinberg S (1994) *Dreams of a Final Theory: The Scientist's Search for the Ultimate Laws of Nature* (Vintage, New York).

144. Trifonov EN (2011) Vocabulary of definitions of life suggests a definition. *J Biomol Struct Dyn* 29:259–266.

145. Koonin EV (2012) Defining life: An exercise in semantics or a route to biological insights? *J Biomol Struct Dyn* 29:603–605.

146. Szostak JW (2012) Attempts to define life do not help to understand the origin of life. *J Biomol Struct Dyn* 29:599–600.

EVOLUTION