INTERFACE FOCUS

royalsocietypublishing.org/journal/rsfs

Review

Cite this article: Solé R, Kempes CP, Corominas-Murtra B, De Domenico M, Kolchinsky A, Lachmann M, Libby E, Saavedra S, Smith E, Wolpert D. 2024 Fundamental constraints to the logic of living systems. *Interface Focus* **14**: 20240010.
https://doi.org/10.1098/rsfs.2024.0010

Received: 11 March 2024
Accepted: 21 August 2024

One contribution of 4 to a theme issue 'Genes, cells and embryos: remembering Pere Alberch'.

Subject Areas:
astrobiology, biocomplexity, systems biology

Keywords:
contingency, convergence, constraints, information, evolution, thermodynamics

Author for correspondence:
Ricard Solé
e-mail: ricard.sole@upf.edu

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

Fundamental constraints to the logic of living systems

Ricard Solé[1,2,3,4], Christopher P. Kempes[4], Bernat Corominas-Murtra[5], Manlio De Domenico[6,7], Artemy Kolchinsky[1,8], Michael Lachmann[4], Eric Libby[4,9], Serguei Saavedra[4,10], Eric Smith[4,11,12] and David Wolpert[4]

[1]ICREA-Complex Systems Lab, Universitat Pompeu Fabra, Dr Aiguader 88, Barcelona 08003, Spain
[2]Institut de Biologia Evolutiva, CSIC-UPF, Pg Maritim de la Barceloneta 37, Barcelona 08003, Spain
[3]European Centre for Living Technology, Sestiere Dorsoduro, 3911, Venezia VE 30123, Italy
[4]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA
[5]Institute of Biology, University of Graz, Graz 8010, Austria
[6]Complex Multilayer Networks Lab, Department of Physics and Astronomy 'Galileo Galilei', University of Padua, Via Marzolo 8, Padova 35131, Italy
[7]Padua Center for Network Medicine, University of Padua, Via Marzolo 8, Padova 35131, Italy
[8]Universal Biology Institute, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[9]Department of Mathematics and Mathematical Statistics, Umeå University, Umeå 90187, Sweden
[10]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA
[11]Department of Biology, Georgia Institute of Technology, Atlanta, GA 30332, USA
[12]Earth-Life Science Institute, Tokyo Institute of Technology, Tokyo 152-8550, Japan

RS, 0000-0001-6974-1008; BC-M, 0000-0001-9806-5643; MDD, 0000-0001-5158-8594; AK, 0000-0002-3518-9208; SS, 0000-0003-1768-363X; DW, 0000-0003-3105-2869

It has been argued that the historical nature of evolution makes it a highly path-dependent process. Under this view, the outcome of evolutionary dynamics could have resulted in organisms with different forms and functions. At the same time, there is ample evidence that convergence and constraints strongly limit the domain of the potential design principles that evolution can achieve. Are these limitations relevant in shaping the fabric of the possible? Here, we argue that fundamental constraints are associated with the logic of living matter. We illustrate this idea by considering the thermodynamic properties of living systems, the linear nature of molecular information, the cellular nature of the building blocks of life, multicellularity and development, the threshold nature of computations in cognitive systems and the discrete nature of the architecture of ecosystems. In all these examples, we present available evidence and suggest potential avenues towards a well-defined theoretical formulation.

1. Introduction

Imagine that a space probe lands on a distant planet. The probe has sophisticated instruments that detect life on different scales. These instruments might detect a network structure of chemical reactions in the atmosphere or identify molecular biosignatures consistent with living systems [1–6]. The probe can also scan its surroundings, capturing morphological biosignatures and other instruments might analyse the chemical network of the atmosphere and measure its topological and molecular complexity [2,7–9].

How different would such an alternative biosphere be? How dependent would an alternative life form be on the environmental context? Are there physical or chemical pre-conditions required for life to emerge? These questions can be extended beyond evolutionary biology [10] and astrobiology [11] and affect our potential to design (using synthetic biology and

THE ROYAL SOCIETY PUBLISHING

© 2024 The Author(s). Published by the Royal Society. All rights reserved.

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

bioengineering) novel life forms [12–15]. More generally, we can capture the essence of the previous questions by asking whether we can predict what kind of (possible) living forms of organization exist beyond what we know from our biosphere (the actual).

One established view sees the evolution of life as highly path-dependent [16–18]. In the words of Jacques Monod [16, pp. 42–43],

> *…the biosphere does not contain a predictable class of objects or events but constitutes a particular occurrence, compatible with first principles but not deducible from these principles, and therefore essentially unpredictable.*

François Jacob, in turn, discussing tinkering in evolution and the problems associated with levels of complexity, concludes that [19]

> *There are always some constraints imposed by stability and thermodynamics. But as complexity increases, additional constraints appear […]. Consequently, there cannot be any general law of evolution.*

Here, Jacob acknowledges the limitations of predicting the properties at one scale based on the properties of the components at lower scales, which is nowadays understood using the concept of *emergent properties* [20,21].

Gould [17] developed the most famous (and controversial) approach to these questions. He argued that because of its historical nature, 're-running the tape of evolution' would lead to entirely different outcomes. Much of the argument was based on the Cambrian Explosion event, which took place 550 Ma ago[1] and involved the rise of all animal body plans [22,23]. Gould's work triggered a renewed interest in the Cambrian event and, more generally, in the problem of evolutionary contingency.

The contingency scenario depicted by Gould has been reanalysed and put in a more general context [24,25]. However, the essential message is still relevant within evolutionary dynamics and astrobiology studies. The evidence for convergent evolution has challenged the idea that an alternative biosphere would look alien to ours [10,26,27]. In contrast to Monod and Gould's views, Conway Morris points out that

> *organisms are under constant scrutiny of natural selection and are also subject to the constraints of the physical and chemical factors that severely limit the action of all inhabitants of the biosphere. Put simply, convergence shows that in the real world, not all things are possible.*

The presence of constraints, particularly in the evolution of developmental programmes, was emphasized in the pioneering work of Pere Alberch, who suggested the concept of 'The logic of monsters': even when dealing with theratologies, whose phenotypic traits have no selective value, we can organize the diversity of forms under a logic taxonomy, suggesting that organismal complexity is strongly limited [28]:

> *…monsters are a good system to study the internal properties of generative rules. They represent forms which lack adaptative function while preserving structural order. There is an internal logic to the genesis and transformation of morphologies and in that logic we may learn about the constraints on the normal.*

In this *structuralist* (or internalist) perspective, gene expression can only be seen as a necessary condition for morphogenetic dynamics, but it is insufficient. As a result of feedback between gene expression patterns and cell–cell interactions, very little of either the structure or the variation in developmental paths is explained by linear mappings from gene labels.[2]

The notion of fundamental constraints is also present in a geometric concept proposed by David Raup: the morphospace, i.e. a multi-dimensional space representing different morphological or structural characteristics of a given class of entities (from cells to networks) [30]. An important lesson from the distribution of living entities across the morphospace is that some parts are densely occupied while others are voids, associated with unobserved possibilities. This uneven occupation is strongly related to the role played by constraints and the presence of convergent solutions. This concept has been extended across disciplines, from network topology to language [31,32], and also applies to ecological systems [28,33]. Like morphospaces, species-interaction networks partition the environmental parameter space into a discrete set of possible biotic configurations, where some partitions can be larger than others [34,35]. This robustness (compatibility with a larger set of environmental conditions) can be a target of evolution, as argued by Waddington and others [36–38].

Finally, there is an argument by Stuart Kauffman and collaborators concerning the intrinsic unpredictability of evolutionary dynamics due to the 'non-ergodic'[3] character of biology [39]. In their own words [40,41]:

> *The chemical and physical properties of the different complex molecules are different, and in biology, the functional properties of these tens of thousands of different molecules in cells are also different. The universe is not ergodic because it will not make all the possible different complex molecules on timescales very much longer than the lifetime of the universe. It is true that most complex things will never 'get to exist'.*

The authors use 'non-ergodic' here in the sense that not every part of the configuration space can be explored. In a nutshell, the size of sequence space Ω associated with a biological polymer of length L built from a molecular alphabet Σ of size $|\Sigma|$ would be

$$\Omega \sim |\Sigma|^L .$$

Thus, the space of possible proteins with a length of 1000 amino acids is 20^{1000}, a space so large that it could never be explored in our universe [42]. The space of possible molecular configurations of molecules within an organism is yet astronomically larger. How can we then talk about universal life features in this scenario?

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

Nonetheless, there are several major reasons to expect convergence. First, all evolutionary trajectories occur under certain generic selective constraints, such as the laws of mathematics, physics and chemistry, and these should lead to some universal features [43,44]. Second, many of these spaces are not explored randomly. For example, analyses of the nature of the genotype spaces show that network structures are far from uniform [45]. This is particularly relevant when dealing with the emergence of molecular functions, where the genotype space is highly redundant [46], meaning that a huge number of genotypes are consistent with the same phenotype [47]. The percolating nature of these genotype spaces strongly favours the potential for success in evolutionary search [48]. Quoting Susanna Manrubia, these properties *further ensure that different functions may await just a few mutations apart* [46,49,50]. Moreover, some intrinsic properties of chemical and physical nature can deeply constrain the possible repertoires of molecular structures, as exemplified by the fact that only a small fraction of protein folds are realizable even looking at the full sequence diversity [51–54]. Finally, evolutionary trajectories happen within a certain system: Darwinian evolution will act on populations of entities. Entities within the system emerge within assembly spaces [2,55,56]. They are constructed from components over evolutionary timescales. These shape evolution and shift the way one should think about sequence search conditioned on the past, where the logic of assembly spaces may lead to certain types of universal convergence [56].

Here, we discuss some of the fundamental constraints that limit the space of evolutionary outcomes. We focus on areas that are most well-studied and most likely to be universal, involving several case studies that reveal deep constraints associated with the *logic* of the organization of living systems. Specifically, we start with some core thermodynamic constraints and then discuss the logic underlying molecular information carriers, cellular reproduction, multicellularity (MC), cognitive architectures and ecosystem organization. Finally, we consider the concept of phase transitions as a paradigm for the emergence of living complexity. A recurrent theme in the following sections concerns predictions about biological complexity made before empirical evidence came in. In our view, such predictions strongly hint at some kind of universality. Such universality fits within the broader goal of finding general theories of life that transcend the specifics of life on Earth [44,57–71].

2. Living systems as thermodynamic engines

Every organism on Earth operates as a thermodynamic engine because it acquires free energy from the environment and uses it to drive essential biological functions. The ubiquitous nature of thermodynamic constraints suggests that a certain kind of *thermodynamic logic* is a universal feature of living systems. In this context, thermodynamics allows us to address the problem of the bounds to the efficiency of living, information-processing agents [72–76].

Any form of life is expected to be an embodied and differentiated structure that performs healing, self-repair and error correction. Such processes reduce the system's entropy by mapping a large set of 'incorrect' (damaged) states to a much smaller set of 'correct' (viable) states. Entropy reduction also comes by growth: the synthesis of organized biological machinery from simpler, disconnected components—for example, as done by the ribosome during the synthesis of proteins—involves a large reduction in entropy [77].

The second law of thermodynamics states that a physical process must generate an overall increase ΔS_{total} of the entropy of a system (ΔS_{system}) and its environment (ΔS_{env}):

$$\Delta S_{\text{total}} = \Delta S_{\text{system}} + \Delta S_{\text{env}} > 0. \tag{2.1}$$

Therefore, living systems can only perform processes that reduce entropy internally if, at the same time, they produce an even greater increase of entropy in the environment. Thus, the universal thermodynamic logic of life is that low-entropy input ('resource') is turned into high-entropy output ('waste'), which can be expressed as[4]

$$\Delta S_{\text{system}} < 0, \quad \Delta S_{\text{env}} > 0, \quad |\Delta S_{\text{env}}| > |\Delta S_{\text{system}}|. \tag{2.2}$$

This suggests that all life forms face the problem of maintaining a low internal entropy in a race against the second law. This thermodynamic perspective on biology is closely associated with Schrödinger's well-known 1944 book *What is life?* [79] and the much less well-known articulation by Boltzmann in 1886 [80].

We may add a few important details and questions to the above-mentioned general picture. First, for a system coupled to a thermal reservoir at temperature T, such as an atmosphere or an ocean, the entropy production in the environment is related to the amount of heat released Q,

$$\Delta S_{\text{env}} = Q/(k_B T), \tag{2.3}$$

where k_B is Boltzmann's constant. In fact, for almost all life on Earth, thermodynamic driving is accomplished by acquiring energy from the environment and releasing it as heat. Important examples of heat generation include the absorption of visible photons and radiation of lower-frequency infrared, the breakdown of high-energy biotically produced molecules (within trophic ecosystems) and the exchange of electrons between environmentally provided donors and acceptors [81]. Although, in principle, entropy production can also occur due to exchanges of other quantities besides energy [82,83], heat flow is a fundamental and ubiquitous energy-exchange channel, suggesting it should have a primary role in the thermodynamics of life on other planets.

Biological heat generation recapitulates the pattern observed at the planetary scale. Late-stage planets are essentially closed to exchanges of matter with their environments but open to energy exchange via incoming solar flux and outgoing thermal radiation [84,85]. Notably, the other source of thermodynamic driving available to chemotrophic life—the gradual oxidation of mantle minerals—needs to be coupled to hydrogen escape to maintain more oxidizing surface conditions than the interior

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

in an era without photosynthetic production of oxidants [86]. However, on a planet with oxygenic photosynthesis, even that process can be maintained while conserving matter, as long as the burial of organic carbon can compensate for the liberation of oxygen—a large-scale geological rearrangement driven by sunlight but mediated by life. To summarize, it is very likely that heat generation serves as a primary thermodynamic driving force for life anywhere. To the extent that life does not completely replace geological processes with novel ones [87], but rather partially conserves geologically facile processes [86], the sharing of major thermodynamic properties by the two becomes even more expected.

A second important point is that the second law is obeyed not only globally—that is, at the level of the entire organisms—but also locally at the level of each individual reaction. For this reason, each entropy-reducing reaction must be locally coupled to a free-energy source.

In biology, a collection of energy intermediates drives many biotically essential transitions that otherwise would not occur spontaneously [88,89]. These intermediates include, first and foremost, the phosphate-bearing cofactors (ATP and the other nucleoside triphosphates, and others [90]) that can drive dehydrating reactions by phosphoryl group transfers, a variety of electron-transfer cofactors (such as NAD, NADP and a variety of others [91]) and membranes that act as capacitors for the exchange of protons. Energy intermediates remove the need for internal processes to be in direct contact with environmental sources of free energy, thus achieving a kind of thermodynamic autonomy [92]. In addition, the fact that many core reactions coupled to these intermediates can be run bidirectionally [93,94] allows life to store energy, e.g. in molecules like glycogen, thereby buffering against stochastic and deterministic (e.g. day/night) environmental fluctuations and achieving further thermodynamic autonomy [92].

As a third point, it is interesting to consider what can be added to the picture using results from non-equilibrium thermodynamics, such as Onsager's principle of detailed balance [95]. One important insight made by Morowitz [96] is that any non-equilibrium chemical system in steady state must exhibit cycles, as illustrated in figure 1. That is, it must exhibit sequences of transformations that leave the system's state invariant while exchanging energy and/or matter with the environment. Today we recognize metabolic cycles, such as the citric acid cycle [98,99], as some of the most fundamental and universal organizing principles of metabolism. Morowitz's insight was that such metabolic cycles must be present in any living system (see figure 1*b*).

To illustrate this concept formally, we may consider a simplified model involving N metabolites, denoted by the vector $\mathbf{C} = (C_1, C_2, \ldots, C_N)$. Suppose that the concentration of each metabolite can be influenced by the reactions converting it from and to other metabolites,

$$\frac{\mathrm{d}C_i}{\mathrm{d}t} = \sum_{j:\, j \neq i} J_{ij}, \tag{2.4}$$

where $J_{ij} = -J_{ji}$ is the net flux of the reaction that converts metabolite j to i. In a steady state, all concentrations are constant, so $\mathrm{d}C_i/\mathrm{d}t = 0$ for all i. The key point is that outside of the trivial (equilibrium) steady state where all fluxes vanish ($J_{ji} = 0$), steady-state fluxes must necessarily exhibit one or more cycles [96,100]. An example is provided by a cycle such as $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$ that corresponds to positive fluxes $J_{21} = J_{32} = J_{13} = 1$. This cycle involves the creation and destruction of metabolites 1,2,3, so its net effect leaves the concentrations invariant.

While living systems must meet the constraints of energy, it should be noted that more proximal constraints from kinetics, from limits of chemical mechanisms or from robust paths for construction and degradation may impose tighter constraints that obscure ultimate roles from thermodynamics. The complexity of identifying the relative importance of ultimate versus proximal constraints is a chief source of difficulty in interpreting based on idealized reversible assembly [101]. However, how key a constraint energy is to living systems can be tested by assessing how close organisms have come to the possibility frontier. Efforts to assess the conversion efficiency in heterotrophs from food to new biomass have been made [102], including a very early effort by Morowitz [103] to tie these explicitly to entropy. More recent work has sought to include the more subtle entropies of sequence information [77] within estimates of efficiency for biological assembly. Other arguments related to metabolic scaling indicate that energy is a key consideration in the evolutionary optimization of organisms [43,44,104,105].

A final note to properly frame the extent of thermodynamic constraints is needed. Many abiotic systems exhibit some of the properties mentioned above. For example, hurricanes, Bénard cells (figure 1*a*) and other naturally occurring 'dissipative structures' [106] also couple local entropy reduction to external entropy production and also exhibit ongoing cycles, yet they are not alive. However, apart from biosynthetic networks in organisms and a few engineered synthetic chemical systems, there are no known dissipative structures whose cycles drive the chemical assembly of complex molecules rather than the mechanical formation of physical structures (such as vortices). Another difference is that abiotic dissipative structures depend entirely on the presence of appropriate boundary conditions for their existence. This is unlike living systems, which construct organismal boundaries and maintain internal energy stores in order to attain a degree of thermodynamic autonomy from immediate environmental conditions.

3. Linear information carriers

Information plays a central role in living systems beyond energy and matter and the thermodynamic considerations described in the previous section. Previous work has suggested that some of the most universal characteristics of life are related to its informational, algorithmic and computational properties [44,62,70,107].

Here, we consider one aspect of this story, motivated by the fact that all lifeforms seem to require physical information carriers that provide the means by which phenotype properties can be transmitted across generations or time. Two properties

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

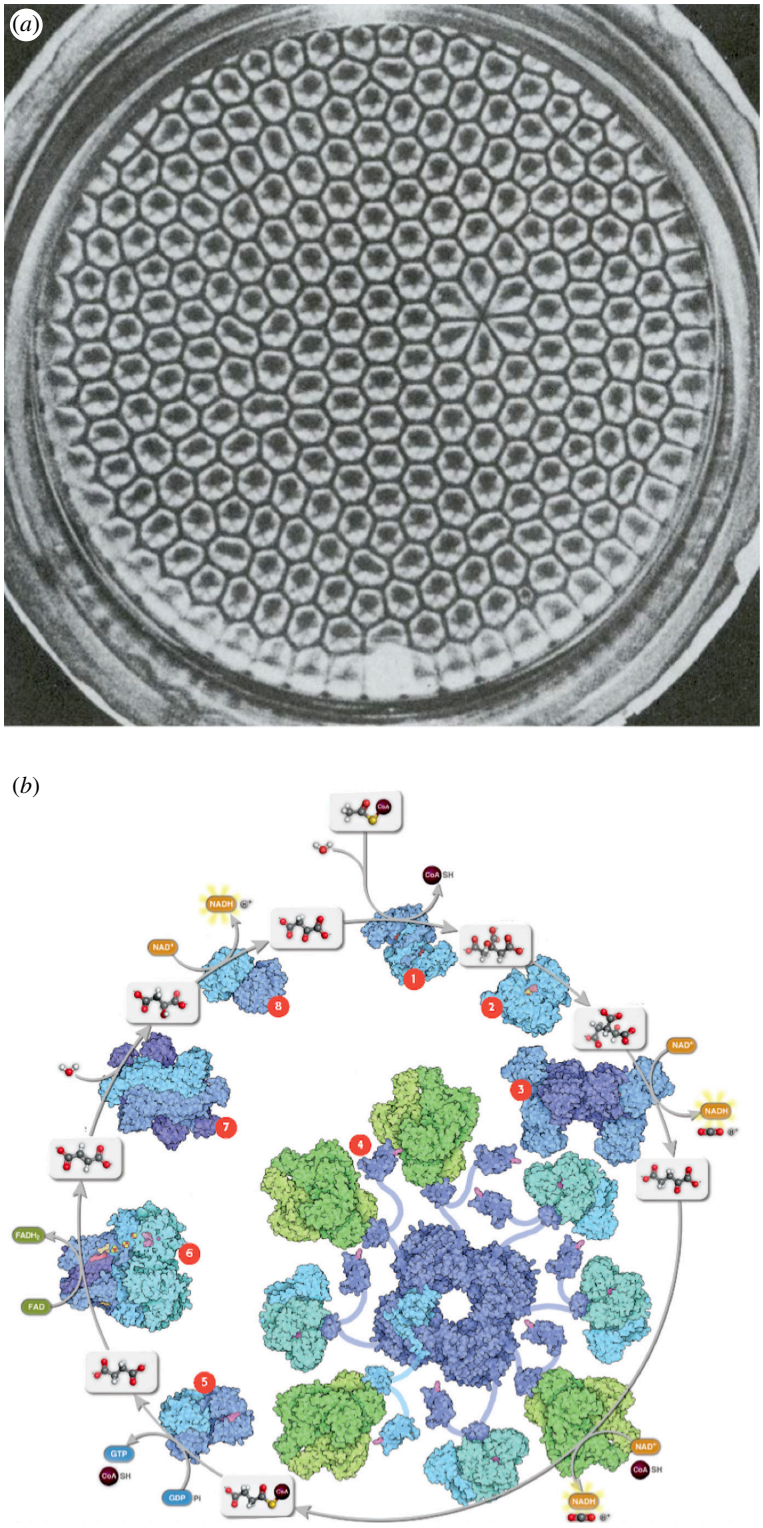


Figure 1. Cyclic structures characterize dissipative systems that reach non-equilibrium steady states due to external driving. In an abiotic system like a Bénard cell (*a*), a temperature gradient leads to the formation of cells that transport heat by cyclic convection (image adapted from Koschmieder & Pallas [97]). In living systems, chemical energy drives metabolic cycles such as the citric acid cycle (*b*), which plays a crucial role in energy production and biosynthesis. Such metabolic cycles consume resource molecules and synthesize energetic intermediates and building blocks while releasing heat. In the figure, both the intermediate metabolites of the cycle and the enzymes are indicated (image from David Goodsell).

of information carriers are important. First, they can reliably code for a large number of phenotypic states. Second, that information can be replicated. Without the emergence of a shared information carrier, any evolved feature also has to evolve a method of transmission of that feature—almost like a new origin of life. What are the fundamental constraints associated with biological information carriers?

Of course, the best-known information carrier in biology is the DNA molecule. Two remarkable insights into the nature of genetic information were advanced before molecular biology and the unveiling of DNA structure in 1953. The first has to be found in the writings of Nikolai Koltsov, who, as early as 1927 concluded that the basis of heritability at the molecular and cellular level had to be found in some class of giant, double-stranded molecules able to self-replicate in a semiconservative

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

way [108,109]. Koltsov's conjecture was that the genetic material replication should be explained in terms of linear copolymers, with each strand to be used as a template [110]. Seventeen years later, another suggestion was found in Schrödinger's book *What is life?*, where he suggested that the information-coding molecule (the gene, pp. 60–61 in [79]) should, on the one hand, be a regular structure (like a 'crystal') while, on the other, allow for an intrinsic 'disorder' compatible with 'atoms playing an individual role'.[5]

Koltsov's and Schrödinger's visions, based on reflections turned out to be strikingly accurate once Watson, Crick and Franklin uncovered the structure of DNA [115,116]. Their work revealed a right-handed helix composed of two antiparallel strands twisted around each other, forming a helical backbone of sugar–phosphate groups with nitrogenous bases paired in the core. As predicted by Koltsov, the discovery elucidated the mechanism of heredity based on semiconservative replication and laid the foundation for understanding how genetic information is stored and transmitted in living organisms. Here, DNA molecules are copied thanks to DNA polymerases, which can 'read' single-stranded DNA chains[6] that act as the 'tape'. This also includes a proofreading mechanism: the DNA polymerase can detect errors in the base pairing and remove the mismatched nucleotides. Moreover, linear tapes and their reading machines are also at work at the levels of transcription (RNA synthesis from DNA; figure 2*a*) and translation (protein synthesis from RNA; figure 2*b*). These systems can be classified as biocomputing machines within a hierarchy [118,119].

Beyond DNA, life, in general, uses three types of polymers: polynucleotides, polypeptides and polysaccharides [120]. Because of their particular chemical features, information and functional (enzymatic and structural) machinery are associated with the first two, while the third group is responsible for energy storage and recognition. The mapping between polymer class and its role is tightly related to their folding, assembly and complementarity properties. Folding allows for the formation of low-entropy states that become stable, thanks to self-interactions, whereas in polynucleotides, monomers do not engage in attractive interactions and display energy degeneracy, i.e. each sequence is energetically about equal. Sugar-based polymers, on the other hand, are branched structures made of the same class of monomers.

The widespread use of long, linear polymers for biological information processing is striking. Is there anything special about one-dimensional linear polymers that leads them to be a universal solution to the information-carrying problem? Are higher-dimensional information carriers possible? Three main arguments support the idea that the linear polymer is the expected option, based on constraints associated with (i) evolvability, (ii) computation and (iii) thermodynamics. Several authors have discussed the first, which concerns the enormous advantages of *some* kinds of linear polymers. As pointed out by Howard Pattee [121] (see also [122–126]):

> *There is an enormously larger class of natural structures that have nearly equal probabilities of formation because they are one-dimensional and have nearly equivalent energies. They are linear copolymers, like polynucleotides and polypeptides. Life and evolution depend on this class of copolymer that forms an unbounded sequence space, undetermined by laws. […] This unbounded sequence space is the first component of the freedom from laws necessary for evolution.*

At the same time, not all information-carrying and information-processing systems are one-dimensional in molecular biology. For instance, the folded DNA molecule exhibits a rich three-dimensional spatial organization, much of which has important functional roles in regulating gene expression. At a more abstract level, cells display a tangled web of molecular interactions connecting different (genetic, metabolic and signalling) networks, which are appropriately conceptualized in high-dimensional space. Nonetheless, much of this complex organization emerges from the capabilities enabled by underlying one-dimensional molecules. This view is partly supported by the observation that higher-dimensional forms of information processing (whether in physical space as with chromatin or abstract space of gene regulatory networks) appear to play a smaller role in more primitive lifeforms, such as prokaryotes.

Beyond the specific solutions found in modern cells, let us consider the following question: what would be the expected logic of a minimal information-processing molecule? This is a relevant problem since any early life capable of evolution should have been able to store and propagate information. Given our current knowledge of molecular cell biology, we might be biased while answering the previous question. In modern cells, linear chains made of discrete units from a given alphabet store information and are scanned by another given molecule that can read the message. However, such a picture was already in place before molecular biology, which Alan Turing introduced in 1936.

The so-called Turing machine is a mathematical model of computation, and its definition will sound very familiar. It consists of an infinite tape divided into cells, a read/write head that moves left or right along the tape and has a finite set of states. Each tape cell contains a symbol from a finite alphabet (figure 2*c*). The machine operates based on a set of transition rules: given the current state and symbol under the head, the machine can write a new symbol, move the head left or right and transition to a new state. The Turing machine represents a foundational concept in the theory of computation, defining a simple computational mechanism capable of performing any algorithmic computation—thereby, it is universal and establishes the basis for understanding the limits and possibilities of algorithmic processes. While building such a general framework, Turing's ideas (perhaps inspired by some of the technology of his time, such as tapes and machines reading them) surprisingly match the molecular computational devices resulting from evolution.

Chemically speaking, some linear polymers seem the simplest and most reliable choices if we need to store bits on a molecular structure. For obvious reasons, a homopolymer (i.e. made of identical monomers) would carry no information. Instead, a heteropolymer, formed by a chain of different kinds of monomers, would provide the substrate of many possible strings of symbols while allowing evolution to occur. Additionally, a molecular system able to scan this string would be strongly constrained by the one-dimensional nature of the heteropolymer. In this context, we could imagine an alternative molecular machinery where a given information substrate is based on some heterogeneous two-dimensional set of monomers and such

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

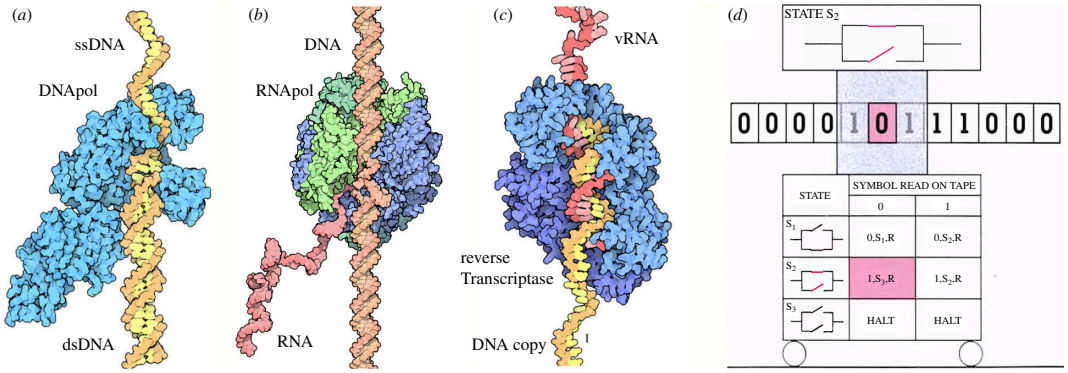


Figure 2. Linear polymers, information and computation. Several molecular events involve information storage and processing, such as DNA replication (*a*) or transcription (*b*) or the RNA → DNA reverse transcription in viruses (*c*). All these examples involve linear polymers that are 'read' by special nanomachines (DNA and RNA polymerases or the reverse transcriptase). Here, RNApol is RNA polymerase, DNApol is DNA polymerase, ssDNA is single-stranded DNA, dsDNA is double-stranded DNA and vRNA is viral RNA. The classical model of computation defined by Turing (*d*) involves a machine with internal states that scans a linear string of symbols (here made of zeros and ones) and changes its internal states as computation proceeds. Images (*a*) and (*b*) are adapted from David Goodsell. Image (*c*) adapted from Hopcroft [117].

potential is illustrated by the remarkable advances in DNA engineering as a material [127,128]. Beyond these nanostructure assembly processes,[7] it has been shown that two-dimensional monolayers can be obtained from mixtures of adenine and uracil [129,130] leading to aperiodic structures, although under contrived physical conditions. In this context, despite the potential information that can be stored in these monolayers and the relevance of surfaces to facilitate polymerization [131], polymers can eventually end up strongly bound to the surface, becoming an evolutionary dead end. Beyond these possibilities, it is not difficult to imagine the challenges imposed by creating, reading and replicating such kind of molecular information in predictable ways.

Our central interest here has been to go beyond the specific to the universal. What can be said of information-processing systems in general? One fundamental constraint is the energy required to perform information-copying operations. Landauer's bound [132] gives us the minimal energy required to perform an abstract computation and many string writing operations, including copies and transformations of it [133]. Copying a DNA strand or producing a protein from a ribosome can be cast as writing a specific string from a set of unordered letters (e.g. nucleotides or amino acids). Under this framework, it is possible to calculate the minimal energy required to write a string, and amazingly, it has been shown that the energy usage of the ribosome is only one or two orders of magnitude above this bound [77].

This is an interesting case where fundamental physical limits are relevant to a very general process, and they also help us to understand cellular physiology. For example, how should we assess whether the energy flux of an environment is sufficient to support a living system? One extreme possibility is to ask if the available energy is sufficient for string copying and processing operations for a very small amount of stored information. One can easily connect the fundamental bounds of information to metabolism via Landauer's bound, by calculating the minimal metabolic rate W needed to replicate genetic information given a genome length and desired cellular growth rate [77]. Landauer's bound states that [132]

$$Q \geq k_B T(S_I - S_F), \tag{3.1}$$

where Q is the heat released to a bath at temperature T, k_B is Boltzmann's constant, S_I and S_F are the initial and final system entropy, respectively. This is nothing more than the expression of the second law of thermodynamics, equations (2.1) and (2.3), where $\Delta S_{\text{sys}} = S_F - S_I$. For the case of writing a specific string from a set of unordered letters, we have that $S_F = 0$ and

$$S_I = \ln\left(n^L\right) = L \ln\, n,$$

n is the number of unique elements, or letters, in the informational system and L is the length of information (the string length) that is being copied. We can convert this into a metabolic rate by considering how fast the information is copied. Given a time to divide t_d in seconds, the minimal metabolic rate of copying is

$$W = k_B T \frac{L \ln n}{t_d}\, \text{J s}^{-1}. \tag{3.2}$$

We illustrate this result in figure 3 for an alphabet of $n = 4$ elements for various L and t_d. This is the minimal metabolic rate for copying alone and does not account for other functions of an organism. We reference the typical bacterial metabolism (TBM) and the typical genome length (TGL) of bacteria. For typical division times, the minimal genome copying cost is many orders of magnitude smaller than known metabolic rates. However, as shown in figure 3, the costs increase with larger genomes and shorter division times.

4. Cells as minimal units of life

All living entities in our current biosphere can be classified as cellular life forms and virus-like elements [135]. Cell division involves, on the one hand, copying the information contained in the cell to its daughter cells, and on the other, reproduction

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

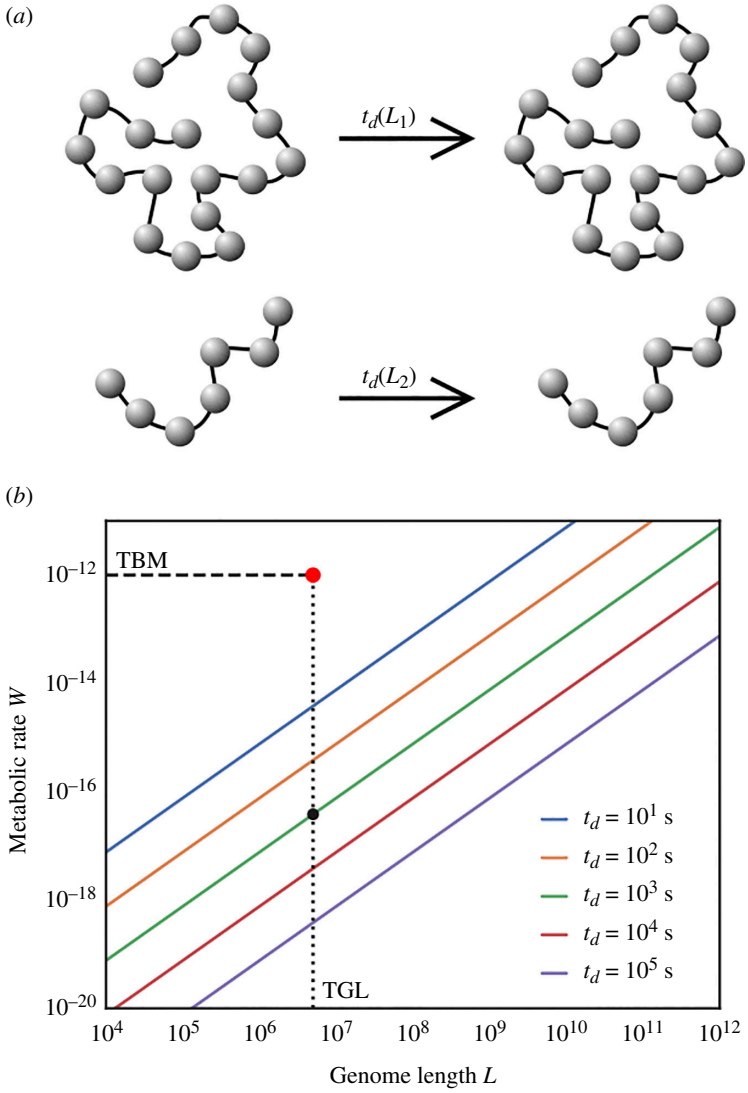


Figure 3. Required metabolic rate for information copying alone, as calculated by equation (3.2). The key parameters are the number of letters in the informational alphabet (n), the length of the genome (L) and the time to copy the information (t_d). We show the minimal metabolic rate for $n = 4$ at various division times t_d, from 1 s to 10^5 s. Dashed lines indicate a TBM rate of $W \approx 10^{-12}$ J s^{-1} [134], TGL of $L \approx 5 \times 10^6$ bases, with typical division time $t_d \approx 10^3$ s.

of its embodied architecture by using this information[8] of its embodied architecture; each copy must define a compartment and a set of metabolic components necessary to start a new cell cycle. More abstractly, both genomes and their vessels require constructive processes. For the genome, the constructive partitioning and assembly occurring in most genotypes' life cycles operate over the more basic copying dynamics, as Watson, Crick and Franklin recognized. The term 'construction dynamics' has been introduced [136,137] to study specific structural factors, ranging from reproduction to ecology and niche construction [138], that place universal constraints or requirements on the dynamics of evolving populations.

In our biosphere, the copying of genomes ultimately depends on the reproduction life cycle bound of metabolizing cells. In that sense, even if they are not singular or even 'minimal' by any unambiguous measure, they are *essential* to the realization of living states on Earth [139–143]. Even the simplest cells exhibit an extraordinary and diverse molecular complexity and common design principles. As far as we know, no alternative building blocks exist in our biosphere. Is the logic of cellular organization an inevitable outcome of the evolution of life? Here, we consider three critical constraints that might impose some fundamental limits to what such a living autonomous agent might be. These include the following three main concepts: (i) the logic of self-replicating machines, (ii) the physicochemical logic of minimal autonomous agents (cells) that become differentiated from their external environment, and (iii) the thermodynamic limitations associated with reliable reproduction.

The algorithmic basis of self-replication was approached by von Neumann using a very abstract (but also general) view, thus ignoring the exact nature of the physical components and the specific functions carried out by the replicator. Von Neumann understood the importance of information and its relevance in providing the instructions necessary to reproduce the entire system while also copying the instructions themselves.[9] To some extent, von Neumann's so-called *Universal Constructor* (UC) was inspired by the steps followed in a factory to build machines in an assembly line, where each component has to be available in space for assembly. Formally, it was defined in terms of a 'machine' that is implemented using operations on a lattice. The machine includes the following four primary components: the Constructor, the Instructions, the Duplicator and the Controller (figure 4*a*). The Constructor (A) builds the new machine out of components from the surrounding environment. The Instructions (I) contain information on how A will operate and effectively define an input tape (as in Turing machines). The Duplicator (B) reads the instructions and duplicates them. Finally, the Controller (C) regulates the whole process, which has to

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

unfold in a given sequence. As defined, the tape plays two markedly different roles. First, the information on the tape provides instructions to be interpreted and allows the construction of a machine. On the other hand, the information on the tape is also treated as uninterpreted data, which must be copied and attached to the new machine.

Von Neumann's insight went a crucial step beyond Schrödinger's conceptualization of information by showing that a self-replicating agent must contain a sufficient description of itself [147,148]. As happened with our previous case study, the components of von Neumann's construction mirror those of self-replication found in cellular biology. Although the biological reality is significantly more complex and multifaceted, we find close similarities between the Duplicator and the information storage mechanism in cells (DNA, perhaps RNA in early protocells), the Controller's role in interpreting and executing instructions resembling cellular control mechanisms, the Constructor's function in manufacturing new components akin to cellular machinery (as executed by RNA polymerase and the ribosome; see figure 4*b*) and the Instructions reflecting the genetic information directing cellular self-replication. These striking similarities suggest a fundamental logic determining the critical components required for a self-replicating system. Once again, von Neumann's theoretical insight came years ahead of discovering the relevant structures in molecular biology. No less important, von Neumann's insight went beyond the problem of self-replication and is deeply ingrained with the problem of open-ended evolution, which has been a central problem within the field of artificial life [149–152] by pointing to the minimal conditions for complexity to be able to grow.

The search for other formal systems able to self-replicate, usually defined on a two-dimensional lattice, has shown simpler examples with a much smaller number of parts than those proposed initially by von Neumann [64,153]. However, a rather crucial problem exists when mapping the original cellular automaton approach to the UC into the real world: all these systems share a high brittleness. Due to the deterministic, spatially dependent nature of the rules required to implement replication, even a slight error (or mutation) typically destroys the whole pattern. Initial conditions must also be fixed in some predetermined way; otherwise, the system will not follow adequate paths towards reliable copying.

How can we solve it? The qualitative theory put in terms of the UC design is agnostic as to how one implements it, and one answer to our question comes from the dominant role played by self-organization of soft matter [154,155] and the requirements to obtain a system displaying agency. Here, chemical and physical constraints play a crucial role. In looking for them, we will also see that alternatives lacking information (a set of instructions) can exist, defining a compartment-metabolism system. More importantly, understanding these phenomena and the potential presence of constraints is essential to understanding the transition from non-living to living matter and the design of artificial cells [142,156–160].

The first constraint for a system to be able to replicate itself is very basic: molecular components need to be available. This is achieved when there is phase separation from the environment [143,161,162] that can concentrate the needed substrates within a given domain. In practical terms, this requires a closed compartment defining an inside and an outside and the subsequent flows of energy and matter that occur across the boundary. Interestingly, a rather limited set of structures are the most plausible candidates. In a water-solvent world, one robust path to creating a compartment that separates the inside from the outside environment is provided by amphiphiles, which spontaneously self-organize in space to form well-defined structures. They are polar molecules with a well-defined hydrophilic head group (attraction to water) and a tail group showing hydrophobicity. Due to this conflicting relation towards water molecules, amphiphiles can self-assemble into compact bilayers that define the system's boundaries. Along with this polar nature, the shape of one molecule largely decides the curvature of the self-organized assembly. The bending energy associated with a given closed configuration is given by

$$\mathcal{H}_b = \oint_S \frac{\kappa[\mathbf{S}]}{2}(C(\mathbf{S}) - C_0(\mathbf{S}))^2 \, \mathrm{d}\mathbf{S},$$

where $\kappa[\mathbf{S}]$ is the bending modulus and $C(\mathbf{S}) - C_0(\mathbf{S})$ is the mean curvature of the vesicle surface at $\mathbf{S}$. Energy minimization, as defined from the solutions of $\delta\mathcal{H}_b = 0$, provides a wide range of possible shapes, from highly stable (spheres, for example, with only positive curvature) to others displaying metastable states that imply the presence of local negative curvature (required when cell division occurs) [163,164].

Although most models of cell origins ignore the physical embodiment defined by interacting amphiphiles, any future development will require this component to explain how evolution allowed the reproduction of early cells. However, physical models already show that there are plenty of opportunities. When membrane growth and permeability are coupled to bending energy, a rich space of great morphological diversity is found [165]. Some toy models also include information (figure 4*c*) coupled to metabolism and compartment growth [166]. But models have also shown that such a possibility exists in an information-free context where only metabolism and compartment are present, both in amphiphile-based vesicles [167–169] and micelles [164,170]. Two examples are displayed in figure 5. In both cases, cell division occurs through growth and instability. The first example, simulated explicitly using dissipative particle dynamics [164], displays no evolution. In the second case, each protocellular assembly carries a different set of molecules, and the division rate depends on the *compositional information* associated with the specific combination of surfactants, so this information indirectly modulates the division process that is driven by the relative free energy difference between the mother aggregate and the resulting daughter aggregates.

None of these information-free examples has been observed in our current biosphere. Why? Although the self-organizing properties of information-free soft matter could provide a source of reliable self-replication, which is not present in von Neumann's formulation, the potential for adaptation and open-ended evolution associated with information-carrying protocells would be difficult to overcome [172]. Thus, we can conjecture that some UC equipped with linear information carriers and exploiting the robustness of self-assembly might have been the expected, perhaps unique solution.

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

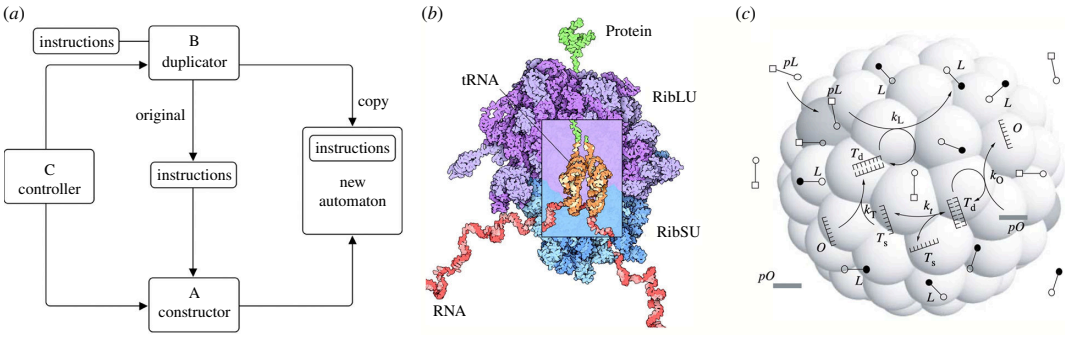


Figure 4. The logic of self-replicating living 'machines'. Cells reproduce through a complex process that uses DNA as a set of instructions but requires also DNA to be replicated. In von Neumann's theory (*a*), a formal machine capable of copying itself would require a set of instructions to guide the construction of a new machine under some controlled states such that instructions also get replicated. In biology, one crucial component of the cellular translation machinery is the ribosome (*b*), made of two subunits (RibLU and RibSU) that 'read' RNA strings to synthesize proteins, playing the role of the Constructor. In a molecular, embodied implementation of cell reproduction, self-organized interactions between a compartment, metabolism and information must interact. An example of a simple implementation is shown in (*c*) for a synthetic cell involving a compartment coupled to double-stranded polymers and a minimal metabolism (adapted from Munteanu *et al*. [146]). Here, a precursor *pL* is transformed into lipids (*L*) that allow membrane growth until some instability triggers division.

5. Multicellularity and development: on growth, form and life cycles

As we mentioned above, the early literature on constraints in development supported the concept that there are limits to the possible in terms of forms and developmental paths [47]. Regarding morphological diversity, multicellular life forms exhibit an enormously rich repertoire of structures. Darwin described this with the famous quote: 'from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved' [173]. For a naturalist, this diversity strikes as the most obvious view of the generative potential of life [174]. But is there a truly endless universe of multicellular form, or is the universe of what is actually observed limited due to fundamental constraints affecting the evolution of complex life?

In this section, we address this question by considering three different problems that include different types of constraints, namely (i) the transitions to individuality within the context of the emergence of MC and the emergence of life cycles; (ii) the nature of the complexity classes in pattern-forming dynamics; and (iii) the existence of a physico-genetic toolkit that pervades the emergence of metazoan complexity.

Evidence suggests that the emergence of cells naturally leads to the potential evolution of MC. MC has emerged independently at least a dozen times on Earth, possibly more, though many lineages have been lost to history [175–177]. Laboratory experiments [178,179] and engineered multicellular systems [180,181] suggest that simple MC[10] is relatively easy to obtain. This is an *evolutionary transition in individuality*. It occurs when replicators, such as cells, form groups that evolve into independent reproducers themselves, causing some loss of autonomy among their parts [182–184] (see also [185–187]). The universal principle is the loss of autonomy in cell-like replicators due to selection acting on groups, favouring traits that enhance group fitness. There is a range of possible MC levels of complexity between strict unicellularity and bona fide multicellular organisms [188]. Crucially, understanding the possible paths to MC requires considering the logic of life cycles [189]. A working definition of MC that encapsulates these features involves the following two properties. (i) Existence. A stage must occur during the organism's life cycle where a group state is clearly recognisable. (ii) Evolution. Groups must be able to multiply and share heritable information with newly created groups [184]. We note that this definition does not explicitly require groups to be formed of the same species, and indeed, there is a vast array of possible multicellular forms involving multiple species [190–192].

What kind of universal logic can be defined here? The first answer to this question comes from the dynamical logic of the problem, which requires the fulfilment of a general principle of biological construction: to create a group, individual units must come together within a finite physical domain and, importantly, deal with the emergence of cheaters [193,194]. In this context, two generic classes of MC can be defined (figure 6*a*,*b*). In the first, MC develops from a single cell C_0 that generates a clonal assembly through cell division ($\mathbf{C} = \{C_k\}$), whereas in the second, there is an aggregation of individual cells from a set $\mathbf{C}_0$ that form a cluster that can be defined by a graph $G = (\mathbf{C}, \omega_{ij})$ defined by the cell locations and their local interactions. These examples illustrate the following two dynamical processes that can generate MC groups: (i) *stay together* (ST) when, as new units are generated, they keep in close connection with the rest, and (ii) *come together* (CT), which occurs when the units move towards each other. Using these basic mechanisms, it is possible to build a taxonomy of cell cycles [189], three of which are indicated in figure 6*b*, as well as mathematical models that allow exploration of the evolutionary principles [198]. These models reveal that ST can favour the division of labour, while CT allows the exploitation of a combination of units with different properties. Both can be found at every level of biological construction, and their dynamical features define constraints to the possible.

There is another general problem associated with MC: the potential for cheaters to emerge. Parasitic entities are a universal property of evolved complexity (see §7), and parasitic entities can threaten the cooperative nature of MC. The problem is illustrated by groups of microbes that produce some common good (such as a metabolite) that helps the group grow faster than others lacking it. However, within groups, free-riding cheats that do not produce the good usually grow fastest of all [199]. The problem arises from the asymmetry in time scales of growth: MC systems evolve slowly compared to the cells within them, which makes them vulnerable to cellular innovations, as occurs with cancer [200]. How can this problem be solved? One solution to stabilizing MC against cheaters

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

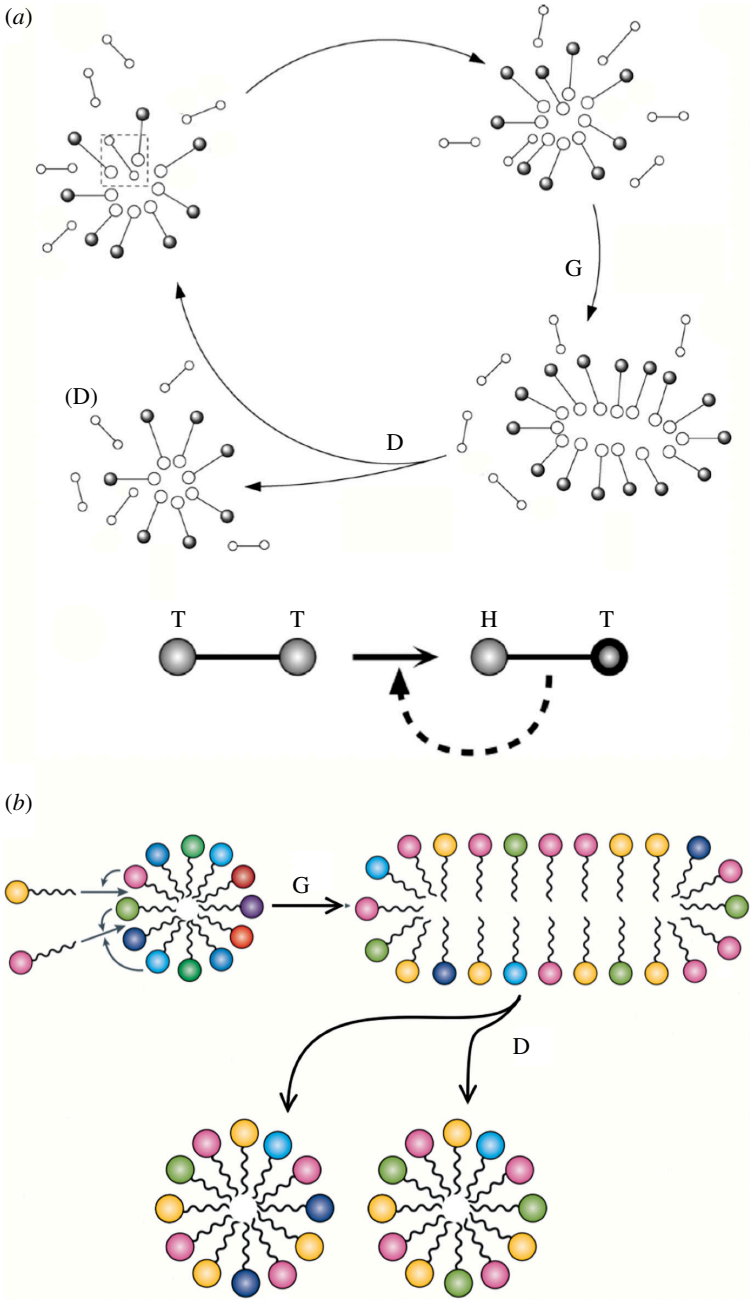


Figure 5. Information-free protocell division. In these two examples, a possible reproduction cycle involving growth G and division D occurs on micelles composed of lipid molecules constantly fed from the environment. In (*a*), these are hydrophobic precursors (indicated as T–T) that react to become a polar amphiphile (H–T), thanks to some reaction (figure adapted from Solé [164]). In (*b*), a similar scenario involves a diverse feeding of lipid molecules forming a heterogeneous aggregate where the recruitment of a particular lipid into the aggregate depends on the compositional information of the aggregate, which lipids already make up the aggregate. In this *lipid world* scenario, the two resulting daughter cells are different (adapted from Kahana & Lancet [171]).

is the evolution of traits that increase cell-level fitness in a group context but come at a cost to free-living fitness. This is enabled by 'ratcheting' processes where cells acquire traits that commit them to a group lifestyle. This stabilizes the group and may serve as a universal mechanism that propels multicellular complexity [201,202]. Developmental ratchets would effectively define an arrow of evolutionary time, allowing increases in developmental complexity.

Early forms of MC are typically quite simple and a far cry from the endless forms that so entranced Darwin. Our last consideration here is tied to pattern-forming mechanisms and the emergence of developmental programmes. The rise of complex animals at the Cambrian boundary brought the question of the creative potential of development and the contributions of chance and necessity. In this context, although many components of the MC toolkit were already in place in the unicellular ancestors [203], black swan events might have also predated the origins of animals [204]. An important consequence of the emergence of gene regulation is the possibility of pattern-forming processes across scales, from cells to organisms [205]. The most famous is the Turing instability associated with reaction–diffusion mechanisms [206,207], which exemplifies the role of self-organization as a mechanism to generate spatial order through symmetry-breaking instabilities [208]. These self-organizing phenomena typically involve the interaction between local amplification (due to reaction dynamics) and long-range communication due to diffusion-like processes. This classic picture has been completed with contributions that incorporate

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

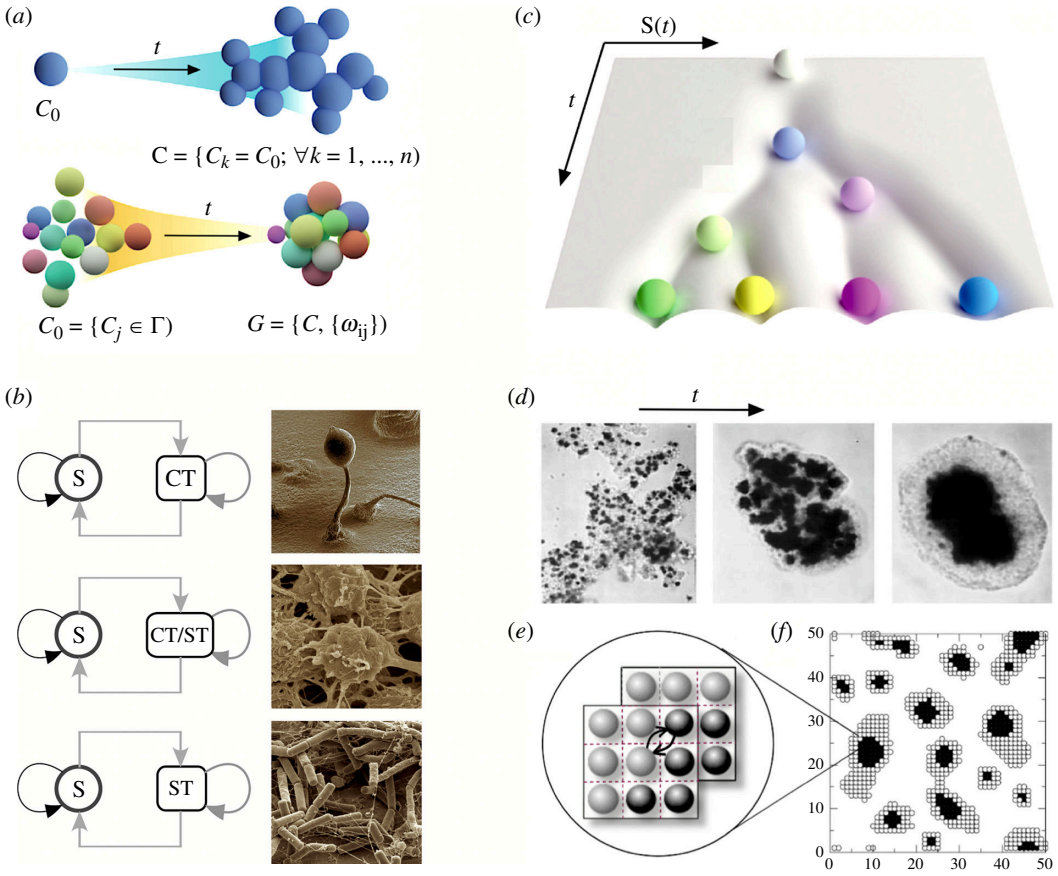


Figure 6. The logic of MC and development. (*a*) At the simplest description level, MC organisms can be assigned to two groups: clonal (upper plot) versus aggregative MC. In the former, a transition $C_0 \rightarrow \mathbf{C}$ occurs after repeated cell divisions with the final population forming a cluster. In the second, the transition is from a set of cells $\mathbf{C}_0$ towards another set (which might contain the same cells and where their states might or might not remain the same) where some interaction matrix ω_{ij} between cell pairs can be defined. (*b*) MC life cycles can be classified within a well-defined taxonomy where transitions to individuality can be described as graphs connecting (grey arrows) single units (*s*) and aggregates that can be generated by coming together (CT) or staying together (ST) mechanisms, or even the coexistence of both. Shown, from top to bottom, are the life cycles of *Dictyostelium discoideum* (image by David Scharf), *Capsaspora owczarzaki* (from the Multigenome Lab) and *Bacillus subtilis* (image by Arnaud Bridier). (*c*) Within MC organisms, cell differentiation increases organismal complexity and can be described by a succession of symmetry-breaking events on a Waddington landscape. Here, marbles indicate cell types (either transient or final). In (*d*) we show an example of the predictable tissue sorting emerging from a completely disordered cell assembly (adapted from Mombach *et al.* [195]). This can be explained through a simple differential adhesion model (*e*,*f*) based on cell sorting dynamics [196]. Here the two cell types are indicated by open and filled circles. Adhesion forces are present, and cells can switch their locations in space if the adhesion energy decreases. Panels (*a*) and (*c*) adapted from Márquez-Zacarías *et al.* [197].

mechanochemical interactions [209,210]. Symmetry breaking is partly responsible for the differentiation paths that allow the generation of specialized cell types [197], often represented as dynamical paths on a Waddington epigenetic landscape (figure 6*c*) [211]. On the other hand, the combinatorial power of gene regulatory networks would also suggest that many (perhaps infinite) potential body plans are possible. To address this problem, one needs to consider the nature of the genotype–phenotype mapping, i.e. the map

$$\Phi : \mathcal{G} \longrightarrow \mathcal{F} \tag{5.1}$$

that can be understood as how a given genotype $\mathbf{G} \in \mathcal{G}$ is connected with a given phenotype $\mathbf{F} \in \mathcal{F}$. The genotype can, for example, be described by the wiring diagram connecting genes, whereas the phenotype could be the observed spatial distribution of cell states. This formal approach has been used in very different contexts to study shape spaces, including RNA folds [212], pattern-forming gene networks [213–216] and circuits [217]. They all share a remarkable universal pattern of organization, including neutrality, characterized by flat regions where different genotypes have the same fitness [29] (see also Catalán *et al*. [218]). They show high redundancy, with many genotypes corresponding to the same phenotype (thus ensuring stability against mutations) and forming interconnected networks. These properties help explain how genetic variation sustains phenotypic stability and drives evolutionary dynamics while, once again, imposing limits to the space of possibilities [214].

The problem of physical constraints to biological form has been addressed historically [219], in particular by D'Arcy Thompson, who proposed the idea that physical analogies between tissues and foams or liquids could help understand the mechanics of biological form and its changes [220,221]. Despite its limitations [222], new theories inspired in physics have validated many of those intuitions [223–225]. In particular, it has been proposed that metazoan complexity can be generated out of a core set of physico-genetic modules affecting well-defined physical properties such as cohesion, viscoelasticity, diffusion or polarity [226,227]. An example is illustrated by the classical work on cell sorting (figure 6*d*), where a randomly mixed cell population displaying two or more cell types experiences a global arrangement due to differential adhesion energies [228,229].

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

As an example, we consider three cell types: $S = 0$ (external medium), $S = 1$ (white) and $S = 2$ (black). Cells attach to neighbours more if it lowers the energy per unit area [230]. The energy at position i, j is

$$\mathcal{H}_{i,j} = \sum_{(k,l) \in \Gamma_{ij}} J_{S_{kl}, S_{ij}}, \tag{5.2}$$

where Γ_{ij} indicates the coordinates of the set of eight nearest neighbours (Moore neighbourhood) and $J_{S_{kl}, S_{ij}}$ indicates the adhesion energy between cell types S_{kl} and S_{ij}.

The dynamics is easily defined: choose one cell located in the coordinate $r = (i, j)$ and see if swapping with a neighbour at $r' = (i', j')$ will or not reduce the energy (thus defining spontaneous transitions). If $\Delta\mathcal{H}$ is the increase of energy in going from the initial to the final state, the probability $P(r' \mid r)$ of swapping is

$$P(r' \mid r) = \frac{1}{1 + e^{\Delta\mathcal{H}/T}}, \tag{5.3}$$

where T is an effective temperature that controls noise. If we iterate this simple probabilistic rule, the system evolves towards a final configuration that can match (in space and time) the observed self-organization towards a stable macroscopic pattern (figure 6*e*,*f*). Single and combinatorial actions of these modules constitute a 'pattern language' capable of generating all metazoan body plans and organs, supporting the view that there are strong bounds to the endless forms.[11]

A final word is in force regarding the collective properties of aggregates of cells: through developmental stages, large-scale but precise deformations are possible, thanks to the collective action of cells [234]. Indeed, even the material properties of single cells may live in the continuum; these individual properties are projected at the collective level—for example, the tissue—in a highly non-trivial way [235,236]. Surprisingly, the continuum of cell material properties is projected to the tissue level as a finite amount of different material phases. This result can be theoretically explained, for example, using a slight variation of the energy function proposed in equation (5.2) [236]. Therefore, MC implies collective behaviour and new ways of regulating biological processes across scales. In §8, we will return to the connection between different scales and how this gives rise to the phenomenon of phase transitions.

6. Cognitive networks, thresholds and brains

Storing and accessing information is necessary for the emergence of cognitive agents [237,238]. However, decision-making and learning were required for higher cognitive complexity and predictive abilities. In other words, living systems evolved mechanisms that could reduce the uncertainty of the environment [239–241]. Moreover, the emergence of multicellular organisms close to the Cambrian explosion precipitated goal-directed movement and behaviours that required enhanced environmental perception and memory [242].

A general feature of many cognitive systems is the presence of mechanisms that transform analogue signals into digital responses. The need for signal discrimination and the fact that analogue computation is more prone to noise might be two crucial constraints on the evolution of cognition [243–246]. Neurons present a prototypical example of digital information processing and its advantages. Although behavioural patterns existed before neurons, the rapid expansion of neural components enabled novel complex behaviours.

The evolution of neurons [247] and neural circuits [248] allowed life to overcome the limitations imposed by diffusion-limited communication [249]. In both neural and aneural agents, all-or-none behavioural decisions are made using a standard design principle: integrating different inputs is weighted, leading to a threshold-mediated response. This also seems to occur within organisms as well as within cells. Is this a universal design principle?

During the mid-twentieth century, a new wave of computing machines created a technological environment for emulating logic elements akin to those in nervous systems. Pioneering theoretical contributions in mathematical biology by Warren McCulloch and Walter Pitts yielded a groundbreaking revelation: the ability to conceptualize units of cognition (referred to as 'neurons') within a logical framework [250,251]. These formal neurons were characterized as threshold units, drawing significant inspiration from the cutting-edge understanding of neural circuits. Unsurprisingly, the McCulloch–Pitts model revolved around the threshold nature of neuron responses and their mathematical depiction [252].

The illustrative depiction of the McCulloch–Pitts model's conceptual framework is outlined in figure 7*a*,*b*. The presented 'formal neuron' (figure 7*b*) operates as a simple Boolean system, with its state assuming one of two values: $S_i \in \Sigma \equiv \{0, 1\}$ (an equivalent description of neurons as spins, $S_i \in \Sigma \equiv \{-1, +1\}$, is sometimes employed when studying neural circuits using techniques from statistical physics). These two states typically correspond to neurons at rest (inactive) or in the act of firing (sending signals to other neurons). In response to incoming signals from a set of N pre-synaptic units, formal neurons exhibit sudden activation if the weighted sum of inputs surpasses a threshold [253,254].

While activation follows an all-or-nothing principle, the weights, denoted as ω_{ij}, are continuous and indicate the impact of the state of pre-synaptic neuron j on post-synaptic neuron i, thereby modelling the strength of connections. These weights can be either positive or negative, allowing for the implementation of excitation and inhibition. In the McCulloch–Pitts approach, the integration of incoming signals by the post-synaptic neuron S_i is expressed as

$$S_i(t+1) = \mathcal{F}\left(\sum_{j=1}^{N} \omega_{ij} S_j(t) - \theta_i\right), \tag{6.1}$$

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

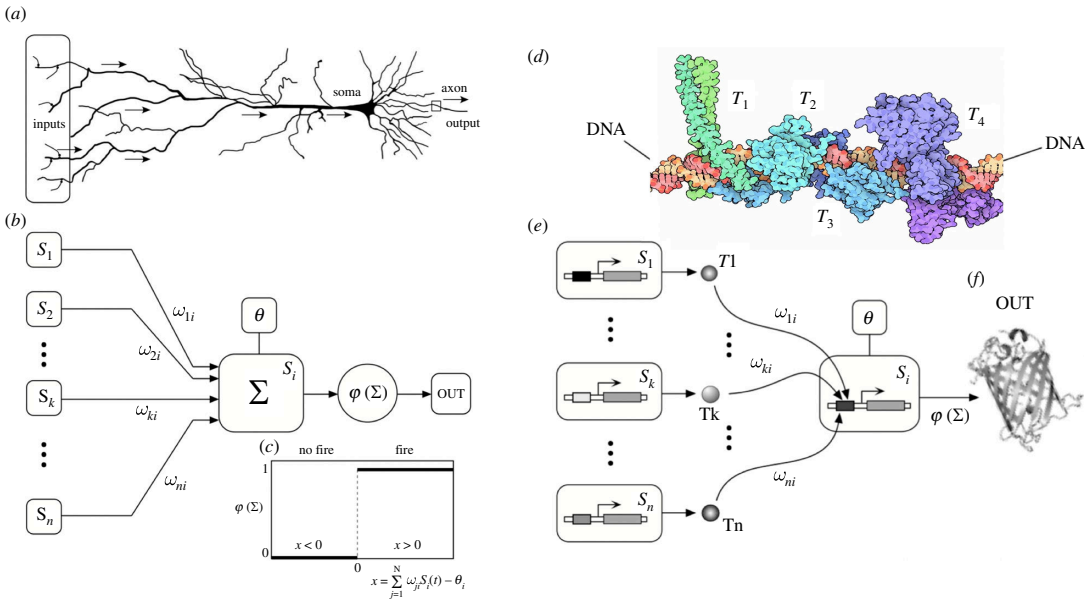


Figure 7. Cognitive networks may exhibit a unique logic of nonlinear response functions. Neurons (*a*) are specialized cells that gather and propagate information in a threshold-like manner. They have a well-defined polarity connected with the input–output signal transmission. The standard McCulloch–Pitts model of a formal neuron (*b*) captures the essence of this transmission in terms of a weighted threshold function, which in its simplest form can be represented as a binary on-or-off response (*c*). Other information transmission systems, like genetic networks, are usually modelled similarly (*d*,*e*). For example, TFs (T_i) are proteins expressed by some genes that bind to DNA (*d*; image by David Goodsell) and regulate the expression of other genes. The input–output diagram in (*e*) is analogous to the neural counterpart, and the corresponding response functions are also highly nonlinear threshold functions. The outcome of these interactions can modify the expression of a given protein (*f*) or set of proteins.

where the additional parameter, θ_i, establishes the neuron's threshold. The nonlinear step function $\mathcal{F}(x)$ yields a value of 1 if its argument is positive and 0 otherwise. Consequently, neuron S_i fires when the weighted sum of pre-synaptic inputs surpasses its threshold. The introduced nonlinearity through $\mathcal{F}(\cdot)$ enforces the all-or-none neural response.[12]

McCulloch and Pitts demonstrated a pivotal insight: combinations of formal threshold neurons can be used to construct any logic Boolean circuit. This implies that, at least in their Boolean representation, brains could execute the same logic operations as computers. The McCulloch–Pitts model and its descendants have influenced the further development of artificial neural networks. Beyond the single-unit design discussed here, artificial neural networks were inspired by another seemingly universal design principle associated with cortical architecture [255]: the presence of multiple processing layers (such as those found in the visual cortex). Are there alternative approaches to designing cognitive networks that do not rely on threshold-like units?

Interestingly, the basic McCulloch–Pitts design is common to many other living systems, including gene regulatory networks [256–259], immune networks [260–262], collective intelligence [263–267] or some aneural systems, such as quorum-sensing decisions in microbial communities, where an explicit equivalence has been defined [268]. These examples share an essential feature that departs from standard neural networks: they are 'liquid', meaning that the parts (proteins, bacteria or ants) move in space and that there is no stable, hardwired connectivity among pairs of individuals [269,270]. Do these systems follow a different logic scheme from the integration-threshold motif?

We illustrate this equivalence by considering gene regulatory networks, one of the best-studied examples. In this case, expression levels of transcription factors (TFs), to be indicated as $[T_i]$ (intracellular concentration) of different TF (T_i in figure 7*d*) change following a dynamical model [271]

$$\frac{\mathrm{d}[T_i]}{\mathrm{d}t} = -\lambda_i[T_i] + \mathcal{G}\left(\sum_{j=1}^{N}\omega_{ij}[T_j](t) - \theta_i\right), \tag{6.2}$$

where λ_i is the degradation rate and $\mathcal{G}$ is a threshold-like function that is a consequence of the order of the molecular nature of interactions between DNA and TF. More precisely, because TF typically forms dimers, the nonlinearity associated with dimerization automatically implies cooperative, threshold-like responses [272,273]. Here, the weights ω_{ij} encapsulate diverse factors influencing the binding of TF.

What are the consequences of employing polarized, threshold-like elements in line with McCulloch–Pitts logic? This approach may facilitate the early development of multilayer processing structures, including interneurons as a crucial innovation, re-entrant closed loops and memory circuits capable of learning [274,275]. Are loops and multilayer architectures a convergent evolutionary design? A positive answer is suggested by the comparative study of neural networks in invertebrates and vertebrates [276,277]. Despite evolving along separate branches, these groups exhibit similar network topologies. This resemblance is observed in layered neural networks described by Ramon y Cajal for insect and cephalopod visual systems

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

[278,279]. It has also been suggested that neural circuits with re-entrant loops are crucial for complex cognitive tasks [280]. One implication of these shared architectures is the potential for convergent minds shaped by evolution [281–284].

7. Ecology: inevitable parasites and functional trees

As the ecologist Ramon Margalef said [285], there is a 'baroque of nature': there are so many species that an inevitable question is 'why so many?' Presumably, an alternative, much simpler prokaryotic biosphere could fulfil all biogeochemical functions of ecosystems [286]. In Margalef's view, even though physical limits bind the organization of ecosystems, there is plenty of room for the possible within these limits. The emergence of variations and the feedback between species and their environments might explain communities' regularities and diversity. And yet, here again, the ecological literature is full of examples of systems where continuous features can be effectively discretized into a small number of categories.

Two basic components of this discretization are involved. First, there is a well-known classification of the interactions between two species into combinations of neutral, positive or negative exchanges [287]. From this set, $\{0, +, -\}$, different pairs can be derived

$$\{\{0,0\}, \{+,+\}, \{-,-\}, \{+,0\}, \{-,0\}, \{+,-\}\}$$

with each pair mapping neutralism, competition, predation or parasitism, among other possibilities.[13] Despite other features affecting each species at play, further classification within each class reveals again a discrete repertoire.[14] Such an approach has been made, for example, within parasites, showing that, despite the enormous diversity of habitats, hosts, sizes or shapes, their diversity of strategies largely transcends phylogenetic boundaries [290].

Are there alternatives to these ecological network architectures? Would ecological webs in other planets (with different energy flows and chemical diversity), display different structural patterns? [291–293]. All these interactions occur (unless under controlled conditions) in a given environmental context where limited resources are available.[15]

Here enters the second component of discretization because the stability properties of these networks largely determine the separation between what is possible and what is not [295,296]. Due to these stability constraints, we can predict that some communities are impossible to observe (cannot be realized). Moreover, despite the enormous variation of climate and resource conditions, ecosystems display essentially the same properties, from Antarctica to the Sahel [297]. We can also reconstruct fossil ecosystems, the so-called 'paleo food webs' [298,299] (figure 8*a*,*b*), and again find universal patterns shared with modern ecological networks [303].

By avoiding the actual embodiment of biological agents and going down to their algorithmic description, *in silico* experiments support the presence of constraints in the logic of possible ecosystems. The first historical attempt to simulate an evolving digital ecology was made in the 1950s by Niles Barricelli, using the MANIAC computer at Los Alamos National Laboratory, as a digital environment [304,305], who wanted to test the theory of *symbiogenesis* [306]. These simulations were based on a grid of cells, each representing an individual organism. The organisms, defined as binary strings, could reproduce, mutate and compete. Complex interactions emerged but with an important characteristic: parasites rapidly evolved, jeopardizing the potential for creating diversity and evolving complexity. As with natural systems, they were a seemingly inevitable component.

Barricelli's ideas were forgotten until the days of artificial life in the 1980s [307–309]. Two crucial developments attracted renewed attention to the evolution of virtual agents: (i) the formalization and propagation of computer viruses and (ii) the creation of computer ecologies. Computer viruses can be seen as an inevitable outcome of computer networks interacting with data storage and transmission processes. Fifty years after Kleene's theorem, self-replicating infectious programmmes were designed to infect local network servers. From a meagre presence in 1990, DOS-based computer viruses skyrocketed to over 10 000 by 1996. This surge prompted an arms race between new viruses and programmers [310], and their spread was shown to follow the epidemic nature of their real counterparts [311]. This is a perfect illustration of a convergent pattern that is shared between living and artificial systems. What about other ecological interactions?

Ecologist Tom Ray designed the program Tierra on a laptop, creating a virtual architecture in the host computer's memory with its instructions and memory space [312,313]. Each digital organism has a genetic code that dictates its behaviour and features. Replication introduces mutation through faulty 'genetic code' copying, while selection happens via competition for memory and CPU resources. Tierra unexpectedly displayed diverse ecological interactions due to the digital selection process. After the growth of shorter, faster-replicating programmes, parasites emerged, needing other programmes for reproduction. Hyper-parasites and immunization mechanisms followed. Recombination (primitive sex) emerged in response to threats, leading to social behaviour through programme cooperation. Other simulations of evolving ecosystems confirm the generative potential of artificial life systems [300,314–317]. Despite differences, the convergence in interaction classes suggests that the discrete set of possible interactions is standard in both digital and natural worlds. Some artificial life models use game-theoretic approaches with coded strategies in a digital genome. Complex patterns emerge, including coevolution, extinction (punctuated equilibrium [318]), cooperation [319] and ecological networks involving trophic levels [301] (figure 8*d*).

These patterns can also be explored using dynamical models that have been used traditionally within the context of artificial chemistries [320,321] and deterministic chemical reaction dynamics [322]. The latter involves systems of coupled differential equations with higher-order nonlinear terms. One standard formulation is

$$\frac{\mathrm{d}N_k}{\mathrm{d}t} = \sum_{i=1}^{s}\sum_{j=1}^{s} \alpha_{ij}^{k} N_j N_i - N_k \Phi(\mathbf{N}), \tag{7.1}$$

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

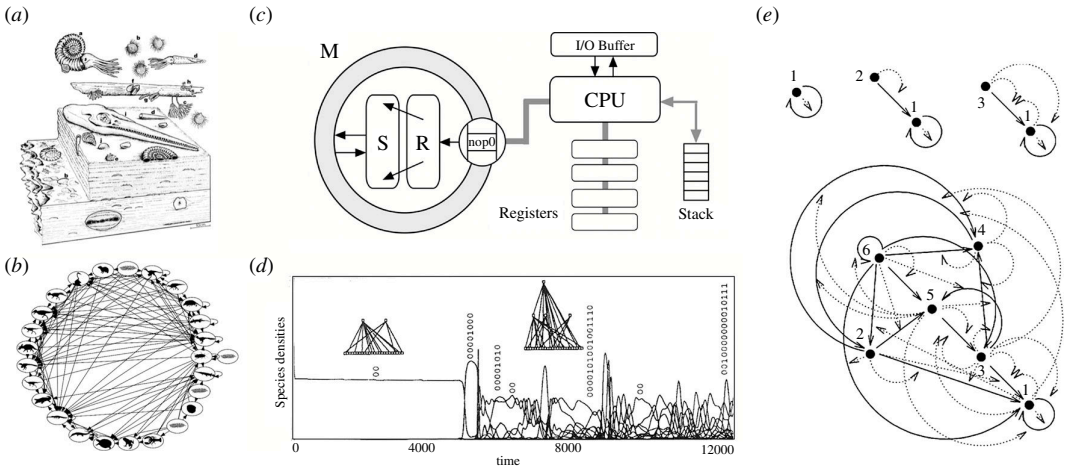


Figure 8. Universal patterns in ecosystems and digital ecologies. Despite their huge diversity, ecosystem architectures are identified in current and fossil communities (*a*) and the ecological network (so-called 'paleo food webs') reconstructed, as shown in (*b*) for a fossil ecosystem before the K–T extinction. One approach to the evolution of these ecological networks relies on digital versions of species and their interactions. An example is shown in (*c*), where the virtual CPU of the Tierra system is summarized (adapted from Adami [300]). Here, S and R stand for the Slicer and Reaper queues, which introduce rewards and ageing. It is possible to evolve food webs (*d*) with a discrete number of layers using evolutionary dynamics on bit strings encoding game-theoretic models (adapted from Kristian & Nordahl [301]). Algorithmic reactors allow evolving interaction networks and see sequences of increase in complexity, as illustrated in (*e*), where different 'species' are indicated as filled nodes, where interactions can happen directly (continuous arrows) or through an intermediate operator rule (adapted from Banzhaf [302]). The evolved networks always include parasitic interactions.

where N_i is the population of type i, $\mathbf{N} = (N_1, ..., N_s)$ is the concentrations vector representing s different types, and each sum includes the different potential reaction events leading to the production of type k, i.e. the bimolecular reactions

$$i + j \longrightarrow i + j + k \tag{7.2}$$

along with their associated rates $\{\alpha_{ij}^k\}$. These are kinetic models grounded in reactions assuming random molecule collisions. The term $\Phi(\mathbf{N})$ stands for a dilution flux that keeps the condition $\sum_j N_j = 1$ of constant population.[16]

How can we introduce evolutionary rules and a context where functionalities are represented? Some particular formulations of these models describe Darwinian dynamics in molecular reaction networks [323] (see also [324]). This occurs when mutations are introduced as one specific reaction class, whereas mean replication is a fitness measure. A finite set of resources provides the source of selection.

The different types of entities whose evolution is described by equations such as equation (7.1) do not need to be restricted to chemical species. An *algorithmic* reaction system can be built using abstract symbols, binary strings, numbers or even proofs [302,321]. In Fontana's *Alchemy* (for algorithmic chemistry), the emergence of novelties is possible by assigning to object interactions the formal properties of functions operating on data [325,326]. In this case, objects act based on symbol binding and λ-calculus rules, where one object functions on another as an argument, resulting in a new object. For two objects $f(x), g(x) \in \mathcal{F}$, an interaction leads to a composed object $f[g(x)]$, creating a new object $h = \phi(f, g) \in \mathcal{F}$. Interactions define a mapping: $\phi : \mathcal{F} \times \mathcal{F} \longrightarrow \mathcal{F}$, representing any computable function, with constraints from underlying semantics. If $\mathcal{P} = 2^{\mathcal{F}}$ (power set of $\mathcal{F}$), an additional mapping $M : \mathcal{P} \longrightarrow \mathcal{P}$ describes the allowable 'collisions'. In the resulting *Turing gas* model, dynamics enable the compositional generation of new objects and the formation of complex interacting object networks, including hypercycles [327–329]. This kind of algorithmic chemistries can lead to complex networks of interactions (starting from catalytic cycles) that also contain parasites (figure 8*e*).

The widespread presence of parasites in both artificial and natural communities suggests that they are a universal outcome of the evolution of complex adaptive systems [330–332]. *In silico* coevolution models [333,334], *in vitro* evolution [335,336] and genomic analyses [337,338] have shown that parasites actively promote diversity and evolvability, as illustrated by the phylogenetic trees in figure 9*a,b*. Here, the presence (figure 9*a*) or absence (figure 9*b*) of parasites dramatically alters species richness, consistently with their role in natural communities [339]. A construct that we term *functional* evolutionary trees can be introduced to provide a unified picture of these case studies. Consider a given evolving biosphere (living or artificial) where novel forms of interactions among agents emerge. Some forms result from adaptations offered by available energy pockets (such as those that occur with parasites), while major innovations mediate others. We could build an evolutionary tree (figure 9*c,d*) where we cluster together those species that share a common set of defining ecological behaviours.

Results from artificial life models, as discussed above, suggest that there is a finite set of possible classes of interactions, namely a set $\Gamma = \{\Gamma_0, \Gamma_1, ..., \Gamma_n\}$, where each Γ_j contains species sharing the same qualitative attributes. Initially, we start from a set of species Σ_0 containing a single homogeneous population of individuals of a single class Γ_0 (such as Ray's initial set of programmes). Due to evolutionary branching, new classes emerge up to a given time where several clusters (classes) Γ_k are found in Σ_t, where the diversity of each class is likely to depend on environmental variables $\{e_j\}$. We conjecture as a result of evolutionary dynamics, the classes form a partition of the final set of species Σ_t, so that

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

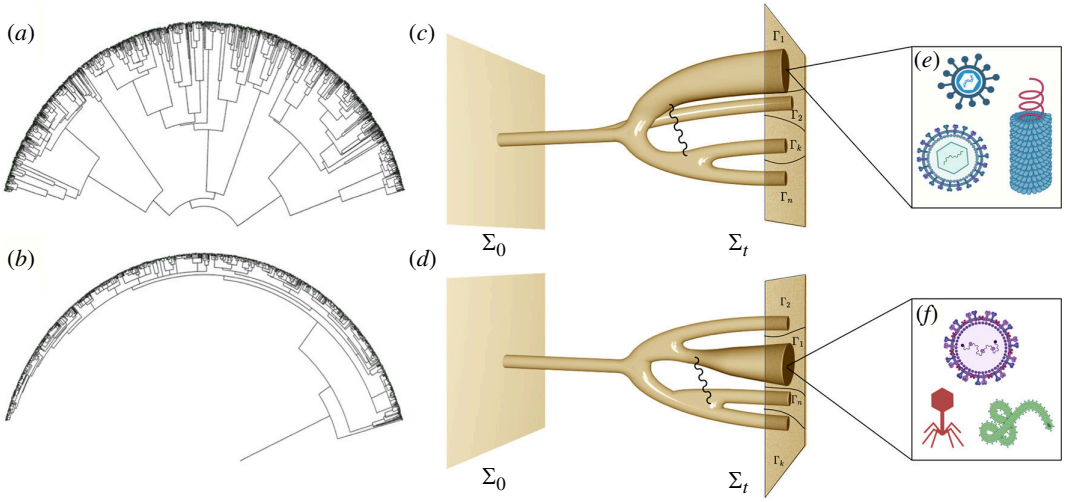


Figure 9. Parasites and functional evolutionary trees. Digital ecologies generate well-defined, qualitative classes of agents that we can identify as parasites, predators or natural counterparts associated with a discrete set of constraints in the repertoire of potential functional roles. In (*a*,*b*), we display the evolutionary trees of an artificial life implementation using *Avida*, where parasites are present (*a*) or suppressed (*b*). The resulting phylogenies indicate that parasites promote diversity (adapted from Zaman *et al*. [333]). The convergent patterns exhibited by ecosystems are illustrated using a diagrammatic representation of two 'runs' (*c*,*d*) of an abstract world, starting from an initial set of species Σ_0 and looking at a final set of species Σ_t. Branches appear each time a new functional class emerges. Although the branching patterns might differ, we conjecture that *functional trees* will branch into the same qualitative, discrete classes $\Sigma_t = \bigcup_{k=1} \Gamma_k$. A thick branch associated with parasites will always be present in both cases, with diverse (but common) solutions indicated on the right (*e*,*f*) using examples from extant viruses (images generated using BioRender). Each time viruses emerge, they will influence the dynamics and even the emergence of new branches, as sketched by the wiggled lines.

$$\Sigma_t = \bigcup_{k=0}^{n} \Gamma_k(e_1^k, \ldots, e_m^k); \quad \Gamma_k \cap \Gamma_j = \varnothing \ \ \forall k \neq j\,. \tag{7.3}$$

Commonalities across different runs would reflect the convergent dynamics of our evolving systems. Specifically, how trees branch over evolutionary time would be somewhat path-dependent (the kinds of parasites might differ; figure 9*e*,*f*). Still, the final structure of ecological networks would include the same set of ecological roles.

Another perspective on ecology follows by considering observed regularities and what they might tell us about life [43,44,340,341]. One of the most powerful recent approaches to ecology connects allometric scaling laws with fundamental physical constraints [43,104,342]. Recently, such approaches to ecology have been extended to create novel biosignatures [340] and to suggest that the logic of cellular metabolism may be universal [341].

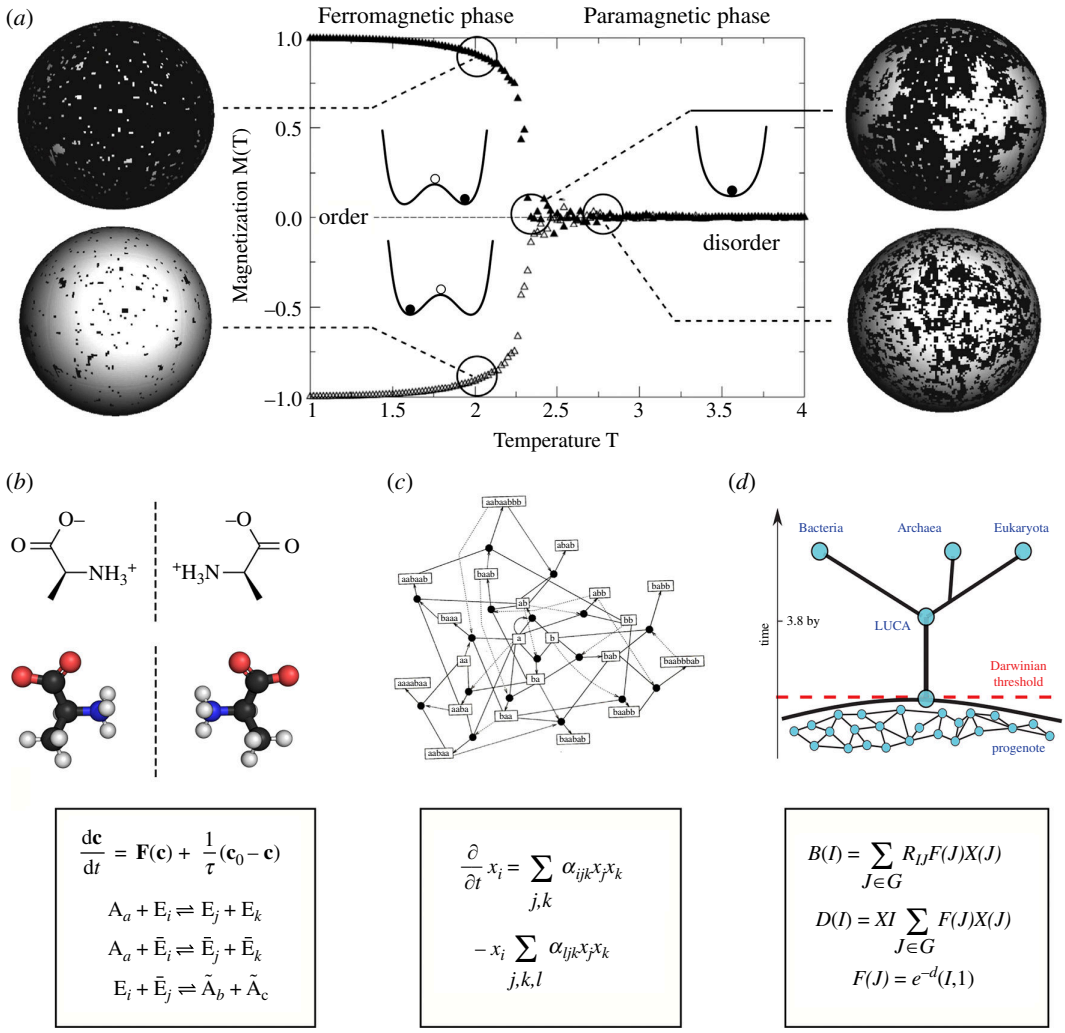
8. Phase transitions and critical states

Some of our previous case studies are deeply connected with the emergence of major innovations in evolution. These so-called *major evolutionary transitions* [343] refer to critical points in the history of life on Earth where new levels of biological organization (cells, MC or language, to cite a few) and complexity emerged. It has been suggested that these transitions, which imply a marked shift from a given qualitative level of organization to a new one, can be mapped into the concept of phase transitions from statistical physics [13,58,61,344,345]. Because of the nature of these transitions, they are likely to be relevant in our understanding of macroevolutionary processes, ranging from punctuated equilibrium [318] to the hierarchical nature of evolutionary change [346,347] and the role that constraints play in setting transition points [105,348,349].

The theory of phase transitions was developed within physics over roughly 90 years spanning most of the twentieth century [350–356], delivering a fully formalized and consistent theory of phases and the transitions separating them. The results have strongly influenced efforts to build a 'systems science' approach to complexity in this area [357,358]. As discussed in §1, the historical nature of living matter seems to play a secondary role when looking at the fundamental logic of its organization. How much of physics constrains the possible, and what is the relative role of contingency? Interestingly, the two components are not exclusive. This is known as *symmetry breaking* [359,360].

We can illustrate this concept using one particularly successful and influential model: the so-called *two-dimensional Ising model* [361–363]. It is a discrete model with simple rules to explain phase transitions in ferromagnetic materials (see figure 10*a*). In this model, a two-dimensional lattice represents a magnet with units S_k ($k = 1, \ldots, N$) having two possible states (spins): up ($S_k = +1$) and down ($S_k = -1$). Quantum mechanics suggests the lowest energy state occurs when adjacent atoms share the same spin. The global magnetization $M = \sum_k S_k$ depends on the difference between the total number of up and down spins (N_+ and N_-). Spins align at low temperatures ($T < T_c$), resulting in ordered states in what is known as the ferromagnetic phase, where $|M| > 0$. Since two possible, completely symmetric ways of alignment are at work, the symmetry must be 'broken'. Above T_c, disorder prevails, with $M \approx 0$. Here, the temperature T is the *control parameter*, whereas the magnetization M is the so-called *order parameter*. As T is tuned, two distinct phases with marked macroscopic properties are observable, separated by a well-defined critical point T_c.

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024



$$\frac{d\mathbf{c}}{dt} = \mathbf{F}(\mathbf{c}) + \frac{1}{\tau}(\mathbf{c}_0 - \mathbf{c})$$

$$A_a + E_i \rightleftharpoons E_j + E_k$$

$$A_a + \bar{E}_i \rightleftharpoons \bar{E}_j + \bar{E}_k$$

$$E_i + \bar{E}_j \rightleftharpoons \tilde{A}_b + \tilde{A}_c$$

$$\frac{\partial}{\partial t} x_i = \sum_{j,k} \alpha_{ijk} x_j x_k - x_i \sum_{j,k,l} \alpha_{ijk} x_j x_k$$

$$B(I) = \sum_{J \in G} R_{IJ} F(J) X(J)$$

$$D(I) = XI \sum_{J \in G} F(J) X(J)$$

$$F(J) = e^{-d}(I,1)$$

Figure 10. Phase transition in physics versus evolutionary transitions. In (*a*), the phase transition associated with the two-dimensional Ising model is shown, displaying the average magnetization (the order parameter) $M(T) = \sum_i S_i$ (i.e. a sum over all spins $S_i \in \{-1, +1\}$) against the temperature (the control parameter). The snapshots are obtained by indicating 'up' (+1) and 'down' (−1) spins as black and white squares, respectively. Two possible ordered configurations are obtained for temperatures lower than the critical T_c, associated with a symmetry-breaking phenomenon. For $T > T_c$, a disordered phase is present, with an essentially random arrangement of spins, lacking spatial correlations. A critical point T_c separates these two phases. The inset curves represent the macroscopic potential function derived from a mean-field approximation. Open and filled circles stand for unstable and stable states, respectively. In (*b*,*c*), we display three examples of models of evolutionary innovations that exhibit some phase transition, along with some of their underlying mathematical descriptions. These are (*b*) the emergence of homochirality (figure from Wikipedia [364]), (*c*) the emergence of autocatalytic cycles out of random chemistry of reacting species (figure redrawn from Farmer *et al*. [365]) and (*d*) the collective evolution of genomes as described by the transition from horizontal to vertical genetic transition (figure from Goldenfeld *et al*. [59]).

The Ising model considers an energy function $\mathcal{H}$, based on the microscopic interactions between magnet units:

$$\mathcal{H}(S_1, \ldots, S_N; J) = -\frac{1}{2}\sum_{\langle i,k \rangle} J S_k S_i. \tag{8.1}$$

The coupling constant J defines the strength of the interaction between nearest neighbours $\langle i, k \rangle$ within the lattice of magnets. Dynamical changes in the orientation of the magnets are introduced using transition probabilities that depend on temperature and neighbour spins. The transition probability is defined using a temperature-dependent function

$$P[S_i \to -S_i] = \frac{1}{1 + e^{\Delta\mathcal{H}/k_B T}}, \tag{8.2}$$

where $\Delta\mathcal{H}$ is the energy change associated with a spin flip.

The Ising model has been shown to predict experimental observations of magnetic systems near $T = T_c$ with high accuracy. This is an important fact since the model ignores most of the microscopic details of the real system, keeping only the symmetry of the interactions, demonstrating universality in critical phenomena theory. Multiple extensions have been proposed, and the number of applications is huge, even considering simple, mean-field versions of the model.[17]

Indeed, beyond ferromagnetism, the two-dimensional Ising model and its extensions have been used in many contexts within complex systems. This includes its equivalence to Eigen's quasispecies model, thus allowing mapping the error threshold as a phase transition [367–369], cell membrane responses [370], multicellular assemblies [371,372], spatiotemporal changes in rainforests [373], universal models of complexity [374] or large-scale functional brain dynamics [375]. We mention this disparate

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

set of examples because, in all cases, it was possible to accurately understand complex phenomena on a qualitative and quantitative basis.

Several well-known examples of evolutionary innovations seem to be associated with a symmetry-breaking event. These include, for example, the transition from 'pre-volution' to evolution [376], natural selection [377], the universality of intermediate metabolism [86,98] as well as the origin of chirality (figure 10*b*) [378–380] as a mechanism to favour one of the two possible (symmetric) solutions through an amplification phenomenon: the final choice would be a historical accident. Other transitions involve the jump to novel properties associated with increased network connectivity. This would be the case of the emergence of autocatalytic cycles out of random chemical reactions (figure 10*c*) once the number of possible reactions crosses the percolation threshold (a non-equilibrium phase transition) [381–384]. Similarly, the collective evolution of the genetic code might have also resulted from a phase transition (figure 10*d*). In this case, the combination of threshold conditions for forming phases and their subsequent robustness in supporting the selection of higher-order organizations was invoked to explain the assignment of amino acids to codons in the genetic code [385]. The insight of this argument, first articulated by Woese in 1967 [386], was that although a *highly reliable* code could function adequately with codon assignments that were a frozen accident as Crick suggested [387], the actual genetic code implemented in molecular machinery is subject to errors even in extant, highly evolved life, and must have been much more error-prone in the earliest eras of ribosomal translation and genome replication and transcription.

Along with the potential connection between evolutionary innovations and phase transitions, critical points play another role. While the previous examples deal with crossing the boundaries from one phase to another, some innovations are tied to the evolution towards critical states. Criticality (sometimes also called 'the edge of chaos') usually refers to the state where the system is poised between two phases: an ordered phase (high order parameter) and a disordered phase (vanishing order parameter). For instance, criticality occurs in the two-dimensional Ising model when $T = T_c$. The idea that criticality might have several desirable properties has been advanced by different scholars within the context of nonlinearity and chaos [388], computation [389], genetic codes [390], virus evolution [391], virus–immune system coevolution [392], neuroscience [393,394] or ecology [395,396]. There are several good reasons for critical states playing a key role in living systems: information transfer becomes optimal at criticality [397,398] and sensitivity to external signals is maximized [399].

One hallmark of these transitions and their macroscopic emergent phases is their significantly lower[18] range of variation compared to the microscopic configurations that produce them. While the possible microscopic configurations can increase combinatorially with system size, the self-reinforcing patterns that emerge can take far fewer forms, sometimes only finitely many, leading to a substantial reduction in the degrees of freedom for the system's potential states [402]. From a developmental perspective—as we already outlined in §5—the existence of material phases at the tissue scale defines a low-dimensional scenario of potential tissue properties and structures with obvious regulatory advantages [403–405]. For example, it has been reported that the embryonic tissue in its early stages lies close to the fluidization critical point, enabling the tissue to *melt* and thus induce a strong deformation to *solidify* again, to fix the morphogenetic changes [406]. This discrete nature comes hand in hand with side effects: cancer progression, in some stages, is more efficient thanks to the fluidization/rigidification of the involved tissues [407]. In addition, huge fluctuations due to the proximity to the critical point may represent a serious drawback for the precision of tissue development [406]. From an evolutionary perspective, these robust properties can be seen differently: phase transitions define a clear boundary separating two qualitative behaviours, with each phase characterized by a few fundamental parameters. Despite their differences, systems that undergo these transitions will exhibit some fundamental common laws of organization. This phenomenon is known as convergence.

One final point is related to the nature of the possible laws that could drive the increasing complexity of living entities. By default, it is assumed that natural selection is the dominant process at work, although alternative scenarios based on persistence-level selection (lacking replication) need some consideration [408]. On the other hand, our previous examples suggest that self-organization and emergent phenomena are as relevant as selection [382]. Is that the case? Could other types of dynamical processes also account for the generation of complexity? Manfred Eigen proposed that natural selection is a phase transition in itself [377], and some authors suggest that this is likely to be a universal principle [409].

This section concludes with an idea suggested by several previous works [61,376,410]: deep constraints also limit the possible kinds of evolutionary laws ruling the biosphere. These would include (along with the thermodynamic laws discussed in §2) (i) an inevitable requirement for autocatalysis [411] as a mechanism for population amplification, (ii) the emergence of molecular heterogeneity as a pre-condition for population dynamics, and (iii) the phase transition to evolution from a non-Darwinian to Darwinian biosphere once some given interaction thresholds are achieved. Future work should consider how a theoretical framework can be defined to prove the uniqueness of natural selection as the expected generative force that drives biological complexity.

9. Discussion

Are there multiple alternative ways to generate living complexity? Predicting potential universal features of living systems, such as those resulting from constraints in the examples discussed in this paper, poses several significant challenges. Given the lack of multiple alternative scenarios, how can we address the logic of life? This question makes particular sense under a view of evolutionary dynamics as a highly path-dependent process. If life (however defined) can unfold in highly dimensional spaces that cannot be fully explored, there seems to be plenty of room for divergent designs. However, the existence of constraints in those spaces of the possible might profoundly change this view. The implications have been discussed in the context of possible life, from chemical constraints to phenotypic convergence in an alternative biosphere. Still, they are likely to be relevant in

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

understanding the limits of bioengineering design. In this context, many of the questions raised in our work have been the target of dedicated efforts within the field of artificial life [300,412–414]. The 'surprising creativity' of digital evolution has been an extraordinary source of inspiration and understanding [415].

In this paper, we conjecture that there are constraints that limit the logic of life across scales. To this goal, we provide a set of arguments suggesting that computational, physical and dynamical constraints profoundly limit the design space of possible living systems. The examples presented here have been supported by different but complementary arguments (computational, ecological, chemical, etc.). Our list (which by no means is intended to be exhaustive) includes the following.

- Internal entropy-reducing processes characterize the thermodynamic logic of living systems. Such processes are enabled by coupling processes that produce greater entropy in the environment, likely in the form of generated heat. Life is also expected to store and employ energy intermediates to drive internal processes and to decouple from environmental conditions, thereby attaining a degree of thermodynamic autonomy. Finally, the internal metabolic process will be organized around cyclic transformations.
- Linear heteropolymers formed by sets of units (symbols) having near-equivalent energies are the expected substrate for carrying molecular information. They allow the exploration of vast combinatorial spaces, and the physical constraints associated with linearity might pose severe limitations to the repertoire of possible monomer candidates.
- Closed cell compartments equipped with a von Neumann replication logic are needed for self-reproducing living forms capable of evolution. The compartment allows the concentration of required molecules and defines a boundary between internal and external environments connected through a membrane that can play a part in the constructor roles by exploiting physical instabilities. Such a closed container can be achieved using a specific class of molecules (the amphiphiles) and is thus constrained to a subset of chemical candidates.
- MC allows the emergence of new kinds of organization out of simpler units. One universal pre-condition for this innovation is the presence of some physically embodied process that guarantees the closeness of cells. While the group provides mechanisms of efficient collective reproduction, these new units of selection (from cell clusters to organisms) need to deal with cheaters through ratcheting. The potential diversity of basic morphological designs might be strongly constrained by a finite number of physico-genetic motifs, whose combinations might generate the whole repertoire of basic developmental programmes sharing deep common morphological motifs.
- Beyond information coding on coded strings, cognitive systems require threshold-like units that allow reliable integration and decision-making. Complex cognition has been unfolding by evolving different (but formally equivalent) circuits based on threshold functions that integrate surrounding signals. In multicellular systems, this means evolving cells that display polarization and provide the means for rapid sensing and propagation of information. Because of these features, complex cognition might have been constrained to evolve towards multilayer systems.
- Ecosystem architectures are deeply constrained within a finite set of possible classes of ecological interactions. Current and past ecosystems reveal such a discrete repertoire of possibilities, and *in silico* models of evolving ecologies support this constrained repertoire. Among other regularities, the widespread presence of parasites suggests that they are an inevitable outcome of complex adaptive systems.

Each particular case questions us about different aspects of the evolutionary logic of biocomplexity. Each of the conjectures requires a rigorous formulation of the hypotheses and, in most cases, the consideration of diverse fields of analysis, from information theory and statistical physics to astrobiology and evolutionary biology. Solutions may require the development of new conceptual frameworks beyond the boundaries associated with each field. Furthermore, the problems under consideration are part of a hierarchy in which certain properties at one scale can affect (or be affected) by those at the next scale.

Does the notion of fundamental constraints for life eliminate the potential for surprises? Certainly not. Concentrating on logical structures ignores an essential aspect of multicellular life: the stunning diversity of morphological, anatomical and physiological adaptations that evolved in response to environmental factors. However, if our proposed constraints are universal, the logic of life elsewhere is likely to be quite familiar.

Ethics. This work did not require ethical approval from a human subject or animal welfare committee.

Data accessibility. This article has no additional data.

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. R.S.: formal analysis, visualization, conceptualization, funding acquisition, investigation, methodology, supervision, validation; C.P.K.: visualization, conceptualization, investigation, supervision, visualization; B.C.-M.: visualization, investigation, supervision; M.D.D.: visualization, investigation, supervision; A.K.: visualization, investigation, supervision, validation; M.L.: visualization, investigation, validation; E.L.: visualization, investigation, validation; S.S.: visualization, investigation, methodology, supervision, validation; E.S.: visualization, formal analysis, investigation, methodology, supervision, validation; D.W.: visualization, investigation, validation.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interests. We declare we have no competing interests.

Funding. M.D.D. acknowledges partial financial support from the Human Frontier Science Program Organization (HFSP Ref. RGY0064/2022). B.C.-M. thanks Adri´an Aguirre-Tamaral for insightful comments on earlier versions of the manuscript and acknowledges the support of the field of excellence 'Complexity of life in basic research and innovation' of the University of Graz and the joint FWF/DFG Weave project I 6533-B. A.K. received funding from the European Union's Horizon 2020 research and innovation programme under Marie Sklodowska-Curie Grant Agreement No. 101068029. E.L. thanks the Swedish Research Council for financial support and the interdisciplinary centres IceLab and the Santa Fe Institute for their creative research environments. E.S. is supported by the Earth-Life Science Institute at the Tokyo Institute of Technology, the Center for the Origin of Life at Georgia Institute of Technology and the Wisconsin Institute for Discovery at the University of

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

Wisconsin at Madison, and thanks the Santa Fe Institute for its generous hospitality during writing. D.W. thanks the Santa Fe Institute for its support. S.S. thanks the support from the National Science Foundation under grant no. DEB-2436069.

Acknowledgements. This paper was facilitated by the coincidental confluence of many of the authors at the Santa Fe Institute (SFI), and we thank the SFI for the general environment that brings such people together. R.S. thanks Luis Seoane, Luis Rocha, Jordi Delgado, Pedro Marquez-Zacarias, Marti Sanchez-Fibla, Clement Moulin-Frier and Kepa Ruiz-Mirazo for comments on the first version of the manuscript, the members of the Complex Systems Lab for so many discussions and the support of an AGAUR FI-SDUR 2020 grant. We thank Thomas Gorochowski for useful comments and Steen Rasmussen for the suggestion of a section on multicellularity, as well as an expanded discussion.

Endnotes

[1]Those estimates are based on evidence from the fossil record that reveals a sharp transition, whereas molecular clock studies (where divergence times are obtained from sequence comparison) locate the origin of animals at about 780 Ma (see [22] and references therein).

[2]The mapping between genes and phenotypes is poorly captured by the linear connection. The nature of the genotype–phenotype mapping [29] is highly compositional and algorithmic, reflecting the inevitable and essential interdependence of the molecular components that constitute organisms and enact their physiology and development.

[3]In statistical physics, an ergodic system is one which, over sufficiently long timescales, explores all possible microstates that are consistent with its macroscopic properties. Mathematically, for any physical observable $\mathcal{A}(\mathbf{q})$, where $\mathbf{q}$ is the microstate that specifies all of the system's coordinates, the long-time average of $\mathcal{A}(\mathbf{q})$ converges to the ensemble average, so

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \mathcal{A}(\mathbf{q}(t))\,\mathrm{d}t = \langle\mathcal{A}\rangle\,.$$

In intuitive terms, an ergodic system explores (over time) all possible ways it can exist. It suggests that, given enough time, the system will visit all the available configurations or states.

[4]In non-equilibrium thermodynamics [78], the rate of entropy production per time τ can be expressed as an integral

$$\Delta S_{\mathrm{tot}}/\tau = \int_V \sigma\,\mathrm{d}V$$

for a system of volume V, with σ indicating the specific entropy production rate, to be written as

$$\sigma = \sum_{k=1}^{n} X_k J_k,$$

where we have a set $\{X_1, ..., X_n\}$ of 'forces' (such as temperature gradients or chemical affinities) and their conjugate fluxes $\{J_1, ..., J_n\}$ (such as flows of heat or reaction rates).

[5]As mentioned above, thermodynamic principles drive the formation of metabolic cycle emergence in living systems, thereby creating interconnected information flow cycles. A generalized information-theoretic framework [111] explains different kinds of structures, using a variational principle to balance diverse information pathways and optimal flow [112]. These network approaches further expand the relevance of constraints at various levels [113,114].

[6]Once helicases unwind the double-stranded DNA helix.

[7]To a large extent, the complementary nature of DNA is what allows enormous flexibility in building structures and machines. A given strand will match only its complement and weakly interact with other sequences. In this way, many components can be assembled together in two or three dimensions while keeping the control on interaction strengths between every pair of elements.

[8]A distinction needs to be made between replicators and reproducers. As suggested by E. Szathmary, reproduction of a cell implies making a copy of the whole that is not limited to the genome, in contrast with viral particles or other putative 'naked' molecular replicators in early life.

[9]Some crucial insights into this problem were advanced within the context of Turing machines. If a universal Turing machine can have access to its own code (a form of self-description), it is possible to show that one can re-write any program so that it will print out a copy of itself before it starts running. In formal terms, for any Turing machine $\mathcal{T}$, there exists a $\mathcal{T}'$ such that $\mathcal{T}'$ prints out a description of $\mathcal{T}'$ on its tape and then behaves in the same way as $\mathcal{T}$. Conceptually, this is reminiscent of the UC, which makes a copy of the Instructions and then carries out the Instructions. The mathematician Stephen Kleene first developed these ideas in his second recursion theorem [144,145].

[10]Simple MC (SMC) refers to a form of multicellular organization where cells form clusters or colonies with limited differentiation and coordination. Unlike complex multicellularity, where cells exhibit specialized functions and are organized into distinct tissues and organs, SMC involves groups of similar or identical cells working together, often with minimal communication or structural integration.

[11]In some cases, as occurs with fractal branching patterns in plants [231] or growing corals [232], there is a set of generative rules (that sometimes can be expressed in terms of grammar) that successfully reproduce the whole repertoire of forms, which can often be classified within a parameter space in terms of separated phases. For intriguing arguments that fractal growth with environmentally regulated parameters governed the Ediacaran fauna that may have been a transition stage to non-self-similar development, see [233].

[12]Alternative implementations may employ smooth step functions, such as

$$\mathcal{F}(x) = \tfrac{1}{1+e^{-\beta x}},$$

where β (an inverse temperature) governs the smoothness, converging to the discontinuous step function as $\beta \to \infty$.

[13]Three of these combinations, $\{+,+\}$, $\{-,-\}$ and $\{+,-\}$, are subdivided when applied to real cases. Predation and parasitism, for example, share the $\{+,-\}$ pair, but the kind of interaction and associated life cycles are different.

[14]It is worth mentioning that there is a branch of theoretical ecology that deals with the *qualitative stability* of communities as derived from the sign matrix of pairwise interactions [288,289]. This indicates that much can be inferred solely based on the qualitative effects of member species on each other.

[15]Importantly, thermodynamic arguments based on the maximum entropy formalism also reveal the power of constraints when dealing with the overall statistical patterns of ecosystem organization (instead of the internal logic) such as species–area relationships or abundance distributions in macroecology [294].

[16]In this case, by making the summation over k we obtain $\Phi(\mathbf{N}) = \sum_{i,j,k}\alpha_{ij}^{k}N_jN_i$ as the associated dilution term.

[17]Ignoring spatial effects, a mean-field approximation can be made for the average magnetization, which follows the differential equation $\frac{\mathrm{d}M}{\mathrm{d}t} = (T_c - T)M - M^3$. It is easy to show that three fixed points exist, associated with a symmetry-breaking (pitchfork) bifurcation [366]. A potential function (as defined in one-dimensional dynamical systems, see [258]) can be derived from $\mathrm{d}M/\mathrm{d}t = -\mathrm{d}V(M)/\mathrm{d}M$, i.e. from

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

$V(M) = -\int \dot{M}\,\mathrm{d}M$, which in this case, gives a fourth-order expression that is a symmetric function, i.e. $V(M) = -(T_c - T)M^2/2 + M^4/4$. For $T > T_c$, the only stable state is $M = 0$, whereas for $T < T_c$, the zero state is unstable, whereas two possible, completely symmetric solutions exist, namely $T_\pm = \pm\sqrt{T_c - T}$. These potential functions are displayed as insets in figure 10*a*.

[18]Formally, this reduction can be described, in the thermodynamic limit, as indefinitely or infinitely lower [400,401].

References

1. Summons RE, Albrecht P, McDonald G, Moldowan JM. 2008 Molecular biosignatures. In *Strategies of life detection* (ed. O Botta), pp. 133–159. New York, NY: Springer. (doi:10.1007/978-0-387-77516-6_11)
2. Marshall SM *et al*. 2021 Identifying molecules as biosignatures with assembly theory and mass spectrometry. *Nat. Commun*. **12**, 3033. (doi:10.1038/s41467-021-23258-x)
3. Chan MA *et al*. 2019 Deciphering biosignatures in planetary contexts. *Astrobiology* **19**, 1075–1102. (doi:10.1089/ast.2018.1903)
4. Seager S, Bains W, Petkowski JJ. 2016 Toward a list of molecules as potential biosignature gases for the search for life on exoplanets and applications to terrestrial biochemistry. *Astrobiology* **16**, 465–485. (doi:10.1089/ast.2015.1404)
5. Johnson SS, Anslyn EV, Graham HV, Mahaffy PR, Ellington AD. 2018 Fingerprinting non-terran biosignatures. *Astrobiology* **18**, 915–922. (doi:10.1089/ast.2017.1712)
6. Cleaves HJ, Hystad G, Prabhu A, Wong ML, Cody GD, Economon S, Hazen RM. 2023 A robust, agnostic molecular biosignature based on machine learning. *Proc. Natl Acad. Sci. USA* **120**, e2307149120. (doi:10.1073/pnas.2307149120)
7. Solé RV, Munteanu A. 2004 The large-scale organization of chemical reaction networks in astrophysics. *Europhys. Lett.* **68**, 170–176. (doi:10.1209/epl/i2004-10241-3)
8. Estrada E. 2012 Returnability as a criterion of disequilibrium in atmospheric reactions networks. *J. Math. Chem*. **50**, 1363–1372. (doi:10.1007/s10910-012-9977-x)
9. Bartlett S, Li J, Gu L, Sinapayen L, Fan S, Natraj V, Jiang JH, Crisp D, Yung YL. 2022 Assessing planetary complexity and potential agnostic biosignatures using epsilon machines. *Nat. Astron.* **6**, 387–392. (doi:10.1038/s41550-021-01559-x)
10. McGhee GR. 2011 *Convergent evolution: limited forms most beautiful*. Cambridge, MA: MIT Press.
11. Grefenstette N *et al*. 2024 Chapter 9: life as we don't know it. *Astrobiology* **24**, S186–S201. (doi:10.1089/ast.2021.0103)
12. Church GM, Regis E. 2014 *Regenesis: how synthetic biology will reinvent nature and ourselves*. New York, NY: Basic Books.
13. Solé R. 2016 Synthetic transitions: towards a new synthesis. *Phil. Trans. R. Soc. B* **371**, 20150438. (doi:10.1098/rstb.2015.0438)
14. Ebrahimkhani MR, Levin M. 2021 Synthetic living machines: a new window on life. *iScience* **24**, 102505. (doi:10.1016/j.isci.2021.102505)
15. Davies J, Levin M. 2023 Synthetic morphology with agential materials. *Nat. Rev. Bioeng*. **1**, 46–59. (doi:10.1038/s44222-022-00001-9)
16. Monod J. 1974 *On chance and necessity*. New York, NY: Harper and Collins.
17. Gould SJ. 1989 *Wonderful life: the burgess shale and the nature of history*. New York, NY: WW Norton & Company.
18. Carroll SB. 2020 *A series of fortunate events: chance and the making of the planet, life, and you*. Princeton, NJ: Princeton University Press. (doi:10.1515/9780691209548)
19. Jacob F. 1977 Evolution and tinkering. *Science* **196**, 1161–1166. (doi:10.1126/science.860134)
20. Anderson PW. 1972 More is different: broken symmetry and the nature of the hierarchical structure of science. *Science* **177**, 393–396. (doi:10.1126/science.177.4047.393)
21. Artime O, De Domenico M. 2022 From the origin of life to pandemics: emergent phenomena in complex systems. *Phil. Trans. R. Soc. A* **380**, 20200410. (doi:10.1098/rsta.2020.0410)
22. Erwin DH, Valentine JW. 2015 Was the Ediacaran–Cambrian radiation a unique evolutionary event? *Paleobiology* **41**, 1–15. (doi:10.1017/pab.2014.2)
23. Carroll SB. 2001 Chance and necessity: the evolution of morphological complexity and diversity. *Nature* **409**, 1102–1109. (doi:10.1038/35059227)
24. Erwin DH. 2015 Was the Ediacaran–Cambrian radiation a unique evolutionary event? *Paleobiology* **41**, 1–15. (doi:10.1017/pab.2014.2)
25. Blount ZD, Lenski RE, Losos JB. 2018 Contingency and determinism in evolution: replaying life's tape. *Science* **362**, eaam5979. (doi:10.1126/science.aam5979)
26. Morris SC. 2003 *Life's solution: inevitable humans in a lonely universe*. Cambridge, UK: Cambridge University Press.
27. Solé RV, Valverde S, Rodriguez-Caso C. 2011 Convergent evolutionary paths in biological and technological networks. *Evol. Educ. Outreach* **4**, 415–426. (doi:10.1007/s12052-011-0346-1)
28. Alberch P. 1989 The logic of monsters: evidence for internal constraint in development and evolution. *Geobios* **22**, 21–57. (doi:10.1016/S0016-6995(89)80006-3)
29. Manrubia S *et al*. 2021 From genotypes to organisms: state-of-the-art and perspectives of a cornerstone in evolutionary dynamics. *Phys. Life Rev*. **38**, 55–106. (doi:10.1016/j.plrev.2021.03.004)
30. Raup DM. 1966 Geometric analysis of shell coiling: general problems. *J. Paleontol.* **40**, 1178–1190.
31. Corominas-Murtra B, Goñi J, Solé RV, Rodríguez-Caso C. 2013 On the origins of hierarchy in complex networks. *Proc. Natl Acad. Sci. USA* **110**, 13316–13321. (doi:10.1073/pnas.1300832110)
32. Avena-Koenigsberger A, Goñi J, Solé R, Sporns O. 2015 Network morphospace. *J. R. Soc. Interface* **12**, 20140881. (doi:10.1098/rsif.2014.0881)
33. Saavedra S, Rohr RP, Bascompte J, Godoy O, Kraft NJB, Levine JM. 2017 A structural approach for understanding multispecies coexistence. *Ecol. Monogr*. **87**, 470–486. (doi:10.1002/ecm.1263)
34. Long C, Deng J, Nguyen J, Liu YY, Alm EJ, Solé R, Saavedra S. 2024 Structured community transitions explain the switching capacity of microbial systems. *Proc. Natl Acad. Sci. USA* **121**, e2312521121. (doi:10.1073/pnas.2312521121)
35. Deng J, Taylor W, Saavedra S. 2022 Understanding the impact of third-party species on pairwise coexistence. *PLoS Comput. Biol*. **18**, e1010630. (doi:10.1371/journal.pcbi.1010630)
36. Waddington CH. 1942 Canalization of development and the inheritance of acquired characters. *Nature* **150**, 563–565. (doi:10.1038/150563a0)
37. Wagner A. 2007 *Robustness and evolvability in living systems*. Princeton, NJ: Princeton University Press.
38. Medeiros LP, Boege K, del-Val E, Zaldívar-Riverón A, Saavedra S. 2021 Observed ecological communities are formed by species combinations that are among the most likely to persist under changing environments. *Am. Nat.* **197**, E17–E29. (doi:10.1086/711663)
39. Kauffman. 2000 *Investigations*. New York, NY: Oxford University Press.
40. Kauffman SA, Roli A. 2023 A third transition in science? *Interface Focus* **13**, 20220063. (doi:10.1098/rsfs.2022.0063)
41. Kauffman SA. 2014 Prolegomenon to patterns in evolution. *Biosystems* **123**, 3–8. (doi:10.1016/j.biosystems.2014.03.004)
42. Dryden DTF, Thomson AR, White JH. 2008 How much of protein sequence space has been explored by life on Earth? *J. R. Soc. Interface* **5**, 953–956. (doi:10.1098/rsif.2008.0085)
43. Kempes CP, Koehl MAR, West GB. 2019 The scales that limit: the physical boundaries of evolution. *Front. Ecol. Evol*. **7**, 242. (doi:10.3389/fevo.2019.00242)

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

44. Kempes CP, Krakauer DC. 2021 The multiple paths to multiple life. *J. Mol. Evol.* **89**, 415–426. (doi:10.1007/s00239-021-10016-2)
45. Aguirre J, Catalán P, Cuesta JA, Manrubia S. 2018 On the networked architecture of genotype spaces and its critical effects on molecular evolution. *Open Biol.* **8**, 180069. (doi:10.1098/rsob.180069)
46. Manrubia S. 2022 The simple emergence of complex molecular function. *Phil. Trans. R. Soc. A* **380**, 20200422. (doi:10.1098/rsta.2020.0422)
47. Alberch P. 1991 From genes to phenotype: dynamical systems and evolvability. *Genetica* **84**, 5–11. (doi:10.1007/BF00123979)
48. Gavrilets S, Gravner J. 1997 Percolation on the fitness hypercube and the evolution of reproductive isolation. *J. Theor. Biol*. **184**, 51–64. (doi:10.1006/jtbi.1996.0242)
49. Fontana W. 2002 Modelling 'evo-devo' with RNA. *Bioessays* **24**, 1164–1177. (doi:10.1002/bies.10190)
50. Schultes EA, Bartel DP. 2000 One sequence, two ribozymes: implications for the emergence of new ribozyme folds. *Science* **289**, 448–452. (doi:10.1126/science.289.5478.448)
51. Li H, Helling R, Tang C, Wingreen N. 1996 Emergence of preferred structures in a simple model of protein folding. *Science* **273**, 666–669. (doi:10.1126/science.273.5275.666)
52. Li H, Tang C, Wingreen NS. 1998 Are protein folds atypical? *Proc. Natl Acad. Sci. USA* **95**, 4987–4990. (doi:10.1073/pnas.95.9.4987)
53. Denton MJ, Marshall CJ, Legge M. 2002 The protein folds as platonic forms: new support for the pre-darwinian conception of evolution by natural law. *J. Theor. Biol*. **219**, 325–342. (doi:10.1006/jtbi.2002.3128)
54. Banavar JR, Maritan A. 2003 Colloquium: geometrical approach to protein folding: a tube picture. *Rev. Mod. Phys.* **75**, 23–34. (doi:10.1103/RevModPhys.75.23)
55. Ahnert SE, Johnston IG, Fink TMA, Doye JPK, Louis AA. 2010 Self-assembly, modularity, and physical complexity. *Phys. Rev. E* **82**, 026117. (doi:10.1103/PhysRevE.82.026117)
56. Sharma A, Czégel D, Lachmann M, Kempes CP, Walker SI, Cronin L. 2023 Assembly theory explains and quantifies selection and evolution. *Nature* **622**, 321–328. (doi:10.1038/s41586-023-06600-9)
57. Cleland CE. 2019 *The quest for a universal theory of life: searching for life as we don't know it*. Cambridge, UK: Cambridge University Press. (doi:10.1017/9781139046893)
58. Goldenfeld N, Woese C. 2011 Life is physics: evolution as a collective phenomenon far from equilibrium. *Annu. Rev. Condens. Matter Phys*. **2**, 375–399. (doi:10.1146/annurev-conmatphys-062910-140509)
59. Goldenfeld N, Biancalani T, Jafarpour F. 2017 Universal biology and the statistical mechanics of early life. *Phil. Trans. R. Soc. A* **375**, 20160341. (doi:10.1098/rsta.2016.0341)
60. Walker SI, Packard N, Cody GD. 2017 Re-conceptualizing the origins of life. *Phil. Trans. R. Soc. A* **375**, 20160337. (doi:10.1098/rsta.2016.0337)
61. Walker SI. 2017 Origins of life: a problem for physics, a key issues review. *Rep. Prog. Phys*. **80**, 092601. (doi:10.1088/1361-6633/aa7804)
62. Davies PCW, Walker SI. 2016 The hidden simplicity of biology. *Rep. Prog. Phys*. **79**, 102601. (doi:10.1088/0034-4885/79/10/102601)
63. Walker SI *et al*. 2018 Exoplanet biosignatures: future directions. *Astrobiology* **18**, 779–824. (doi:10.1089/ast.2017.1738)
64. Langton CG. 1984 Self-reproduction in cellular automata. *Physica D* **10**, 135–144. (doi:10.1016/0167-2789(84)90256-2)
65. von Neumann J, Burks A (eds). 1966 *Theory of self-reproducing automata*. Champaign, IL: University of Illinois Press.
66. Langton CG, Taylor C, Farmer JD, Rasmussen S. 1992 *Artificial life II*. Redwood City, CA: Addison-Wesley.
67. Langton CG, Taylor C, Farmer JD, Rasmussen S. 1994 *Artificial life III*. Redwood City, CA: Addison-Wesley.
68. Küppers BO. 1990 *Information and the origin of life*. Heidelberg, Germany: Springer.
69. Yockey HP. 2005 *Information theory, evolution, and the origin of life*. Cambridge, UK: Cambridge University Press. (doi:10.1017/CBO9780511546433)
70. Walker SI, Davies PC. 2013 The algorithmic origins of life. *J. R. Soc. Interface* **10**, 20120869. (doi:10.1098/rsif.2012.0869)
71. Flack J. 2017 Life's information hierarchy. In *From matter to life: information and causality* (eds SI Walker, PC Davies, GF Ellis), pp. 283–302. Cambridge, UK: Cambridge University Press. (doi:10.1017/9781316584200.012)
72. Bennett CH. 1982 The thermodynamics of computation—a review. *Int. J. Theor. Phys.* **21**, 905–940. (doi:10.1007/BF02084158)
73. Bennett CH, Landauer R. 1985 The fundamental physical limits of computation. *Sci. Am.* **253**, 48–57. (doi:10.1038/scientificamerican0785-48)
74. Wolpert DH. 2016 The free energy requirements of biological organisms; implications for evolution. *Entropy* **18**, 138. (doi:10.3390/e18040138)
75. Kolchinsky A, Wolpert DH. 2018 Semantic information, autonomous agency and non-equilibrium statistical physics. *Interface Focus* **8**, 20180041. (doi:10.1098/rsfs.2018.0041)
76. Piñero J, Solé R, Kolchinsky A. 2024 Optimization of nonequilibrium free energy harvesting illustrated on bacteriorhodopsin. *Phys. Rev. Res.* **6**, 013275. (doi:10.1103/PhysRevResearch.6.013275)
77. Kempes CP, Wolpert D, Cohen Z, Pérez-Mercader J. 2017 The thermodynamic efficiency of computations made in cells across the range of life. *Phil. Trans. R. Soc. A* **375**, 20160343. (doi:10.1098/rsta.2016.0343)
78. Glansdorff P, Prigogine I. 1964 On a general evolution criterion in macroscopic physics. *Physica* **30**, 351–374. (doi:10.1016/0031-8914(64)90009-6)
79. Schrödinger E. 1944 *What is life? The physical aspect of the living cell and mind*. London, UK: Cambridge University Press.
80. Boltzmann L. 1974 The second law of thermodynamics. In *Theoretical physics and philosophical problems*, pp. 13–32. Dordrecht, The Netherlands: Springer. (doi:10.1007/978-94-010-2091-6_2)
81. Lengeler JW, Drews G, Schlegel HG. 1999 *Biology of the prokaryotes*. NY: Blackwell Science. (doi:10.1002/9781444313314)
82. Liu J -S., Marison IW, von Stockar U. 2001 Microbial growth by a net heat up-take: a calorimetric and thermodynamic study on acetotrophic methanogenesis by *Methanosarcina barkeri* . *Biotechnol. Bioeng*. **75**, 170–180. (doi:10.1002/bit.1176)
83. Heijnen JJ, Van Dijken JP. 1992 In search of a thermodynamic description of biomass yields for the chemotrophic growth of microorganisms. *Biotechnol. Bioeng*. **39**, 833–858. (doi:10.1002/bit.260390806)
84. Barlow C, Volk T. 1990 Open systems living in a closed biosphere: a new paradox for the gaia debate. *BioSystems* **23**, 371–384. (doi:10.1016/0303-2647(90)90018-v)
85. Budyko MI. 2012 *The evolution of the biosphere*. Berlin, Germany: Springer Science & Business Media.
86. Smith E, Morowitz HJ. 2016 *The origin and nature of life on Earth: the emergence of the fourth geosphere*. London, UK: Cambridge University Press.
87. Benner SA, Ellington AD, Tauer A. 1989 Modern metabolism as a palimpsest of the RNA world. *Proc. Natl Acad. Sci. USA* **18**, 7054–7058. (doi:10.1073/pnas.86.18.7054)
88. Branscomb E, Russell MJ. 2013 Turnstiles and bifurcators: the disequilibrium converting engines that put metabolism on the road. *Biochim. Biophys. Acta* **1827**, 62–78. (doi:10.1016/j.bbabio.2012.10.003)
89. Branscomb E, Biancalani T, Goldenfeld N, Russell M. 2017 Escapement mechanisms and the conversion of disequilibria; the engines of creation. *Phys. Rep.* **677**, 1–60. (doi:10.1016/j.physrep.2017.02.001)
90. Pasek MA, Kee TP. 2011 On the origin of phosphorylated biomolecules. In *Origins of life: the primal self-organization* (eds R Egel, DH Lankenau, AY Mulkidjanian), pp. 57–84. Berlin, Germany: Springer-Verlag. (doi:10.1007/978-3-642-21625-1_3)
91. Braakman R, Smith E. 2013 The compositional and evolutionary logic of metabolism. *Phys. Biol*. **10**, 011001. (doi:10.1088/1478-3975/10/1/011001)
92. Bergareche AM, Ruiz-Mirazo K. 1999 Metabolism and the problem of its universalization. *BioSystems* **49**, 45–61. (doi:10.1016/S0303-2647(98)00034-3)
93. Braakman R, Smith E. 2012 The emergence and early evolution of biological carbon-fixation. *PLoS Comput. Biol*. **8**, e1002455. (doi:10.1371/journal.pcbi.1002455)

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

94. Nunoura T *et al*. 2018 A primordial and reversible TCA cycle in a facultatively chemolithoautotrophic thermophile. *Science* **359**, 559–563. (doi:10.1126/science.aao3407)
95. Onsager L. 1931 Reciprocal relations in irreversible processes. I. *Phys. Rev*. **37**, 405–426. (doi:10.1103/PhysRev.37.405)
96. Morowitz HJ. 1966 Physical background of cycles in biological systems. *J. Theor. Biol*. **13**, 60–62. (doi:10.1016/0022-5193(66)90007-5)
97. Koschmieder EL, Pallas SG. 1974 Heat transfer through a shallow, horizontal convecting fluid layer. *Int. J. Heat Mass Transf*. **17**, 991–1002. (doi:10.1016/0017-9310(74)90181-1)
98. Smith E, Morowitz HJ. 2004 Universality in intermediary metabolism. *Proc. Natl Acad. Sci. USA* **101**, 13168–13173. (doi:10.1073/pnas.0404922101)
99. Morowitz H, Smith E. 2007 Energy flow and the organization of life. *Complexity* **13**, 51–59. (doi:10.1002/cplx.20191)
100. Schnakenberg J. 1976 Network theory of microscopic and macroscopic behavior of master equation systems. *Rev. Mod. Phys*. **48**, 571–585. (doi:10.1103/RevModPhys.48.571)
101. England JL. 2013 Statistical physics of self-replication. *J. Chem. Phys.* **139**, 121923. (doi:10.1063/1.4818538)
102. Calow P. 1977 Conversion efficiencies in heterotrophic organisms. *Biol. Rev*. **52**, 385–409. (doi:10.1111/j.1469-185X.1977.tb00840.x)
103. Morowitz HJ. 1955 Some order-disorder considerations in living systems. *Bull. Math. Biophys*. **17**, 81–86. (doi:10.1007/BF02477985)
104. West GB, Brown JH, Enquist BJ. 1997 A general model for the origin of allometric scaling laws in biology. *Science* **276**, 122–126. (doi:10.1126/science.276.5309.122)
105. Kempes CP, Dutkiewicz S, Follows MJ. 2012 Growth, metabolic partitioning, and the size of microorganisms. *Proc. Natl Acad. Sci. USA* **109**, 495–500. (doi:10.1073/pnas.1115585109)
106. Prigogine I, Nicolis G. 1971 Biological order, structure and instabilities. *Q. Rev. Biophys.* **4**, 107–148. (doi:10.1017/s0033583500000615)
107. Cronin L *et al*. 2006 The imitation game—a computational chemical approach to recognizing life. *Nat. Biotechnol.* **24**, 1203–1206. (doi:10.1038/nbt1006-1203)
108. Koltsov N. 1927 Physical-chemical fundamentals of morphology. *Prog. Exp. Biol.* **7**, 3–31.
109. Koltsov N. 1935 Hereditary molecules. *Sci. Life* **5**, 4–314.
110. Soyfer VN. 2001 The consequences of political dictatorship for Russian science. *Nat. Rev. Genet.* **2**, 723–729. (doi:10.1038/35088598)
111. De Domenico M, Biamonte J. 2016 Spectral entropies as information-theoretic tools for complex network comparison. *Phys. Rev. X* **6**, 041062. (doi:10.1103/PhysRevX.6.041062)
112. Ghavasieh A, De Domenico M. 2024 Diversity of information pathways drives sparsity in real-world networks. *Nat. Phys*. **20**, 512–519. (doi:10.1038/s41567-023-02330-x)
113. De Domenico M. 2023 More is different in real-world multilayer networks. *Nat. Phys*. **19**, 1247–1262. (doi:10.1038/s41567-023-02132-1)
114. Artime O, Grassia M, De Domenico M, Gleeson JP, Makse HA, Mangioni G, Perc M, Radicchi F. 2024 Robustness and resilience of complex networks. *Nat. Rev. Phys*. **6**, 114–131. (doi:10.1038/s42254-023-00676-y)
115. Watson JD, Crick FH. 1953 Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**, 737–738. (doi:10.1038/171737a0)
116. Franklin RE, Gosling RG. 1953 Evidence for 2-chain helix in crystalline structure of sodium deoxyribonucleate. *Nature* **172**, 156–157. (doi:10.1038/172156a0)
117. Hopcroft JE. 1984 Turing machines. *Sci. Am*. **250**, 86–98. (doi:10.1038/scientificamerican0584-86)
118. Amos M. 2004 *Cellular computing*. Oxford, UK: Oxford University Press.
119. Al-Hashimi HM. 2023 Turing, von Neumann, and the computational architecture of biological machines. *Proc. Natl Acad. Sci. USA* **120**, e2220022120. (doi:10.1073/pnas.2220022120)
120. Runnels CM, Lanier KA, Williams JK, Bowman JC, Petrov AS, Hud NV, Williams LD. 2018 Folding, assembly, and persistence: the essential nature and origins of biopolymers. *J. Mol. Evol*. **86**, 598–610. (doi:10.1007/s00239-018-9876-2)
121. Pattee H. 1969 How does a molecule become a message? communication in development. In *Developmental biology supplement* (ed. L Anton). New York, NY: Academic Press. (doi:10.1016/B978-0-12-395541-8.50006-9)
122. Kull K. 2002 A sign is not alive—a text is. *Σημειωτκή—Sign Syst. Stud.* **30**, 327–336. (doi:10.12697/SSS.2002.30.1.21)
123. Barbieri M. 2003 *The organic codes: an introduction to semantic biology*. Cambridge, UK: Cambridge University Press. (doi:10.1017/CBO9780511614019)
124. Pattee HH. 2008 Physical and functional conditions for symbols, codes, and languages. *Biosemiotics* **1**, 147–168. (doi:10.1007/s12304-008-9012-6)
125. Rocha LM, Hordijk W. 2005 Material representations: from the genetic code to the evolution of cellular automata. *Artif. Life.* **11**, 189–214. (doi:10.1162/1064546053278964)
126. Goodwin B. 2017 Biology and meaning. In *Towards a theoretical biology*, vol. 4, pp. 259–275. Edinburgh, UK: Edinburgh University Press.
127. Seeman NC. 2003 DNA in a material world. *Nature* **421**, 427–431. (doi:10.1038/nature01406)
128. Ramezani H, Dietz H. 2020 Building machines with DNA molecules. *Nat. Rev. Genet*. **21**, 5–26. (doi:10.1038/s41576-019-0175-6)
129. Sowerby SJ, Stockwell PA, Heckl WM, Petersen GB. 2000 Self-programmable, self-assembling two-dimensional genetic matter. *Orig. Life Evol. Biosph*. **30**, 81–99. (doi:10.1023/a:1006616725062)
130. Sowerby SJ, Holm NG, Petersen GB. 2001 Origins of life: a route to nanotechnology. *BioSystems* **61**, 69–78. (doi:10.1016/s0303-2647(01)00130-7)
131. Cairns-Smith AG. 1988 The chemistry of materials for artificial Darwinian systems. *Int. Rev. Phys. Chem*. **7**, 209–250. (doi:10.1080/01442358809353213)
132. Landauer R. 1961 Irreversibility and heat generation in the computing process. *IBM J. Res. Dev.* **5**, 183–191. (doi:10.1147/rd.53.0183)
133. Kolchinsky A, Corominas-Murtra B. 2020 Decomposing information into copying versus transformation. *J. R. Soc. Interface* **17**, 20190623. (doi:10.1098/rsif.2019.0623)
134. Milo R, Phillips R. 2015 *Cell biology by the numbers*. New York, NY: Garland Science.
135. Koonin EV. 2014 The origins of cellular life. *Antonie Van Leeuwenhoek* **106**, 27–41. (doi:10.1007/s10482-014-0169-5)
136. Flack JC. 2012 Multiple time-scales and the developmental dynamics of social systems. *Phil. Trans. R. Soc. B* **367**, 1802–1810. (doi:10.1098/rstb.2011.0214)
137. Flack JC, Erwin D, Elliot T, Krakauer DC. 2013 Timescales, symmetry, and uncertainty reduction in the origins of hierarchy in biological systems. In *Evolution cooperation and complexity* (eds K Sterelny, R Joyce, B Calcott, B Fraser), pp. 45–74. Cambridge, MA: MIT Press.
138. Odling-Smee FJ, Laland KN, Feldman MW. 2003 *Niche construction: the neglected process in evolution*. Princeton, NJ: Princeton Univesity Press.
139. Oró J, Miller SL, Lazcano A. 1990 The origin and early evolution of life on Earth. *Annu. Rev. Earth Planet. Sci*. **18**, 317–356. (doi:10.1146/annurev.ea.18.050190.001533)
140. Vernadsky VI. 1998 *The biosphere*. New York, NY: Springer Science & Business Media.
141. Peretó J. 2012 Out of fuzzy chemistry: from prebiotic chemistry to metabolic networks. *Chem. Soc. Rev*. **41**, 5394–5403. (doi:10.1039/c2cs35054h)
142. Ruiz-Mirazo K, Briones C, de la Escosura A. 2014 Prebiotic systems chemistry: new perspectives for the origins of life. *Chem. Rev*. **114**, 285–366. (doi:10.1021/cr2004844)
143. Deamer DW. 1997 The first living systems: a bioenergetic perspective. *Microbiol. Mol. Biol. Rev*. **61**, 239–261. (doi:10.1128/mmbr.61.2.239-261.1997)
144. Kleene SC. 1938 On notation for ordinal numbers. *J. symb. log*. **3**, 150–155. (doi:10.2307/2267778)
145. Marion JY. 2012 From Turing machines to computer viruses. *Phil. Trans. R. Soc. A* **370**, 3319–3339. (doi:10.1098/rsta.2011.0332)
146. Munteanu A, Attolini CSO, Rasmussen S, Ziock H, Solé RV. 2007 Generic Darwinian selection in catalytic protocell assemblies. *Phil. Trans. R. Soc. Lond. B* **362**, 1847–1855. (doi:10.1098/rstb.2007.2077)
147. Pattee HH. 1977 Dynamic and linguistic modes of complex systems. *Int. J. Gen. Syst*. **3**, 259–266. (doi:10.1080/03081077708934771)
148. Rocha LM. 2001 Evolution with material symbol systems. *BioSystems* **60**, 95–121. (doi:10.1016/s0303-2647(01)00110-1)

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

149. McMullin B. 2000 John von Neumann and the evolutionary growth of complexity: looking backward, looking forward. *Artif. Life* **6**, 347–361. (doi:10.1162/106454600300103674)
150. Ruiz-Mirazo K, Umerez J, Moreno A. 2008 Enabling conditions for 'open-ended evolution'. *Biol. Philos.* **23**, 67–85. (doi:10.1007/s10539-007-9076-8)
151. Ruiz-Mirazo K, Peretó J, Moreno A. 2004 A universal definition of life: autonomy and open-ended evolution. *Orig. Life Evol. Biosph.* **34**, 323–346. (doi:10.1023/b:orig.0000016440.53346.dc)
152. Packard N, Bedau MA, Channon A, Ikegami T, Rasmussen S, Stanley KO, Taylor T. 2019 An overview of open-ended evolution: editorial introduction to the open-ended evolution II special issue. *Artif. Life* **25**, 93–103. (doi:10.1162/artl_a_00291)
153. Sipper M. 1998 Fifty years of research on self-replication: an overview. *Artif. Life* **4**, 237–257. (doi:10.1162/106454698568576)
154. Mouritsen OG. 2005 *Life—as a matter of fat*. Berlin, Germany: Springer.
155. Hamley IW. 2007 *Introduction to soft matter: synthetic and biological self-assembling materials*. London, UK: John Wiley & Sons.
156. Rasmussen S, Chen L, Deamer D, Krakauer DC, Packard NH, Stadler PF, Bedau MA. 2004 Evolution. Transitions from nonliving to living matter. *Science* **303**, 963–965. (doi:10.1126/science.1093669)
157. Stano P, Luisi PL. 2010 Achievements and open questions in the self-reproduction of vesicles and synthetic minimal cells. *Chem. Commun.* **46**, 3639–3653. (doi:10.1039/b913997d)
158. Luisi PL. 2016 *The emergence of life: from chemical origins to synthetic biology*. Cambridge, UK: Cambridge University Press.
159. Serra R *et al*. 2017 *Modelling protocells*. New York, NY: Springer.
160. Rasmussen S, Bedau MA, McCaskill JS, Packard NH. 2009 A roadmap to protocells. In *Protocells: bridging nonliving and living matter*, pp. 71–100. Cambridge, MA: MIT Press. (doi:10.7551/mitpress/9780262182683.003.0004)
161. Morowitz HJ, Heinz B, Deamer DW. 1988 The chemical logic of a minimum protocell. *Orig. Life Evol. Biosph.* **18**, 281–287. (doi:10.1007/BF01804674)
162. Deamer D. 2011 *First life: discovering the connections between stars, cells, and how life began*. Berkeley, CA: University of California Press.
163. Solé RV, Munteanu A, Rodriguez-Caso C, Macía J. 2007 Synthetic protocell biology: from reproduction to computation. *Phil. Trans. R. Soc. B* **362**, 1727–1739. (doi:10.1098/rstb.2007.2065)
164. Solé RV. 2009 Evolution and self-assembly of protocells. *Int. J. Biochem. Cell Biol.* **41**, 274–284. (doi:10.1016/j.biocel.2008.10.004)
165. Ruiz-Herrero T, Fai TG, Mahadevan L. 2019 Dynamics of growth and form in prebiotic vesicles. *Phys. Rev. Lett*. **123**, 038102. (doi:10.1103/PhysRevLett.123.038102)
166. Fellermann H, Rasmussen S, Ziock HJ, Solé RV. 2007 Life cycle of a minimal protocell—a dissipative particle dynamics study. *Artif. Life* **13**, 319–345. (doi:10.1162/artl.2007.13.4.319)
167. Macía J, Solé RV. 2007 Protocell self-reproduction in a spatially extended metabolism-vesicle system. *J. Theor. Biol*. **245**, 400–410. (doi:10.1016/j.jtbi.2006.10.021)
168. Macía J, Solé RV. 2007 Synthetic Turing protocells: vesicle self-reproduction through symmetry-breaking instabilities. *Phil. Trans. R. Soc. B* **362**, 1821–1829. (doi:10.1098/rstb.2007.2074)
169. Corominas-Murtra B. 2019 Thermodynamics of duplication thresholds in synthetic protocell systems. *Life* **9**, 9. (doi:10.3390/life9010009)
170. Fellermann H, Solé RV. 2007 Minimal model of self-replicating nanocells: a physically embodied information-free scenario. *Phil. Trans. R. Soc. B* **362**, 1803–1811. (doi:10.1098/rstb.2007.2072)
171. Kahana A, Lancet D. 2021 Self-reproducing catalytic micelles as nanoscopic protocell precursors. *Nat. Rev. Chem.* **5**, 870–878. (doi:10.1038/s41570-021-00329-7)
172. Vasas V, Szathmáry E, Santos M. 2010 Lack of evolvability in self-sustaining autocatalytic networks constraints metabolism-first scenarios for the origin of life. *Proc. Natl Acad. Sci. USA* **107**, 1470–1475. (doi:10.1073/pnas.0912628107)
173. Darwin C. 1859 *On the origin of species*. London, UK: John Murray.
174. Wilson EO. 1992 *The diversity of life*, 1st edn. Cambridge, MA: Harvard University Press.
175. Grosberg RK, Strathmann RR. 2007 The evolution of multicellularity: a minor major transition? *Annu. Rev. Ecol. Evol. Syst.* **38**, 621–654. (doi:10.1146/annurev.ecolsys.36.102403.114735)
176. Bonner JT. 2009 *First signals: the evolution of multicellular development*. Princeton, NJ: Princeton University Press.
177. Knoll AH. 2011 The multiple origins of complex multicellularity. *Annu. Rev. Earth Planet. Sci.* **39**, 217–239. (doi:10.1146/annurev.earth.031208.100209)
178. Ratcliff WC, Denison RF, Borrello M, Travisano M. 2012 Experimental evolution of multicellularity. *Proc. Natl Acad. Sci. USA* **109**, 1595–1600. (doi:10.1073/pnas.1115323109)
179. Ratcliff WC, Fankhauser JD, Rogers DW, Greig D, Travisano M. 2015 Origins of multicellular evolvability in snowflake yeast. *Nat. Commun.* **6**, 6102. (doi:10.1038/ncomms7102)
180. Solé R, Ollé-Vila A, Vidiella B, Duran-Nebreda S, Conde-Pueyo N. 2018 The road to synthetic multicellularity. *Curr. Opin. Syst. Biol.* **7**, 60–67. (doi:10.1016/j.coisb.2017.11.007)
181. Toda S, Blauch LR, Tang SK, Morsut L, Lim WA. 2018 Programming self-organizing multicellular structures with synthetic cell-cell signaling. *Science* **361**, 156–162. (doi:10.1126/science.aat0271)
182. Michod RE. 2007 Evolution of individuality during the transition from unicellular to multicellular life. *Proc. Natl Acad. Sci. USA* **104**, 8613–8618. (doi:10.1073/pnas.0701489104)
183. West SA, Fisher RM, Gardner A, Kiers ET. 2015 Major evolutionary transitions in individuality. *Proc. Natl Acad. Sci. USA* **112**, 10112–10119. (doi:10.1073/pnas.1421402112)
184. Libby E, B Rainey P. 2013 A conceptual framework for the evolutionary origins of multicellularity. *Phys. Biol*. **10**, 035001. (doi:10.1088/1478-3975/10/3/035001)
185. Carmel Y. 2023 Human societal development: is it an evolutionary transition in individuality? *Phil. Trans. R. Soc. B* **378**, 20210409. (doi:10.1098/rstb.2021.0409)
186. Rainey PB. 2023 Major evolutionary transitions in individuality between humans and AI. *Phil. Trans. R. Soc. B* **378**, 20210408. (doi:10.1098/rstb.2021.0408)
187. Krakauer D, Bertschinger N, Olbrich E, Flack JC, Ay N. 2020 The information theory of individuality. *Theory Biosci*. **139**, 209–223. (doi:10.1007/s12064-020-00313-7)
188. Queller DC, Strassmann JE. 2009 Beyond society: the evolution of organismality. *Phil. Trans. R. Soc. B* **364**, 3143–3155. (doi:10.1098/rstb.2009.0095)
189. van Gestel J, Tarnita CE. 2017 On the origin of biological construction, with a focus on multicellularity. *Proc. Natl Acad. Sci. USA* **114**, 11018–11026. (doi:10.1073/pnas.1704631114)
190. Queller DC. 2000 Relatedness and the fraternal major transitions. *Phil. Trans. R. Soc. Lond. B* **355**, 1647–1655. (doi:10.1098/rstb.2000.0727)
191. Libby E, Ratcliff WC. 2021 Lichens and microbial syntrophies offer models for an interdependent route to multicellularity. *Lichenologist* **53**, 283–290. (doi:10.1017/S0024282921000256)
192. Andersson R, Isaksson H, Libby E. 2022 Multi-species multicellular life cycles. In *The evolution of multicellularity*, pp. 343–356. Boca Raton, FL: CRC Press. (doi:10.1201/9780429351907-21)
193. Aktipis CA, Boddy AM, Jansen G, Hibner U, Hochberg ME, Maley CC, Wilkinson GS. 2015 Cancer across the tree of life: cooperation and cheating in multicellularity. *Phil. Trans. R. Soc. B* **370**, 20140219. (doi:10.1098/rstb.2014.0219)
194. Márquez-Zacarías P, Conlin PL, Tong K, Pentz JT, Ratcliff WC. 2021 Why have aggregative multicellular organisms stayed simple? *Curr. Genet*. **67**, 871–876. (doi:10.1007/s00294-021-01193-0)
195. Mombach JC, Glazier JA, Raphael RC, Zajac M. 1995 Quantitative comparison between differential adhesion models and cell sorting in the presence and absence of fluctuations. *Phys. Rev. Lett.* **75**, 2244–2247. (doi:10.1103/PhysRevLett.75.2244)

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

196. Ollé-Vila A, Duran-Nebreda S, Conde-Pueyo N, Montañez R, Solé R. 2016 A morphospace for synthetic organs and organoids: the possible and the actual. *Integr. Biol.* **8**, 485–503. (doi:10.1039/c5ib00324e)
197. Márquez-Zacarías P, Pineau RM, Gomez M, Veliz-Cuba A, Murrugarra D, Ratcliff WC, Niklas KJ. 2021 Evolution of cellular differentiation: from hypotheses to models. *Trends Ecol. Evol.* **36**, 49–60. (doi:10.1016/j.tree.2020.07.013)
198. Tarnita CE, Taubes CH, Nowak MA. 2013 Evolutionary construction by staying together and coming together. *J. Theor. Biol*. **320**, 10–22. (doi:10.1016/j.jtbi.2012.11.022)
199. Sandoz KM, Mitzimberg SM, Schuster M. 2007 Social cheating in *Pseudomonas aeruginosa* quorum sensing . *Proc. Natl Acad. Sci. USA* **104**, 15876–15881. (doi:10.1073/pnas.0705653104)
200. Aktipis A, Maley CC. 2017 Cooperation and cheating as innovation: insights from cellular societies. *Phil. Trans. R. Soc. B* **372**, 20160421. (doi:10.1098/rstb.2016.0421)
201. Libby E, Conlin PL, Kerr B, Ratcliff WC. 2016 Stabilizing multicellularity through ratcheting. *Phil. Trans. R. Soc. B* **371**, 20150444. (doi:10.1098/rstb.2015.0444)
202. Libby E, Ratcliff WC. 2014 Ratcheting the evolution of multicellularity. *Science* **346**, 426–427. (doi:10.1126/science.1262053)
203. Sebé-Pedrós A, Degnan BM, Ruiz-Trillo I. 2017 The origin of metazoa: a unicellular perspective. *Nat. Rev. Genet*. **18**, 498–512. (doi:10.1038/nrg.2017.21)
204. Ruiz-Trillo I, Kin K, Casacuberta E. 2023 The origin of metazoan multicellularity: a potential microbial black swan event. *Annu. Rev. Microbiol*. **77**, 499–516. (doi:10.1146/annurev-micro-032421-120023)
205. Murray JD. 2003 *Mathematical biology: II. Spatial models and biomedical applications*. New York, NY: Springer. (doi:10.1007/b98869)
206. Turing AM. 1990 The chemical basis of morphogenesis. *Bull. Math. Biol.* **52**, 153–197. (doi:10.1007/BF02459572)
207. Marcon L, Sharpe J. 2012 Turing patterns in development: what about the horse part? *Curr. Opin. Genet. Dev.* **22**, 578–584. (doi:10.1016/j.gde.2012.11.013)
208. Prigogine I, Nicolis G. 1967 On symmetry-breaking instabilities in dissipative systems. *J. Chem. Phys.* **46**, 3542–3550. (doi:10.1063/1.1841255)
209. Odell G, Oster G, Burnside B, Alberch P. 1980 A mechanical model for epithelial morphogenesis. *J. Math. Biol.* **9**, 291–295. (doi:10.1007/BF00276030)
210. Collinet C, Lecuit T. 2021 Programmed and self-organized flow of information during morphogenesis. *Nat. Rev. Mol. Cell Biol.* **22**, 245–265. (doi:10.1038/s41580-020-00318-6)
211. Waddington Ch. 2014 *The strategy of the genes*. London, UK: Routledge.
212. Fontana W, Schuster P. 1998 Shaping space: the possible and the attainable in RNA genotype-phenotype mapping. *J. Theor. Biol.* **194**, 491–515. (doi:10.1006/jtbi.1998.0771)
213. Solé RV, Fernández P, Kauffman SA. 2003 Adaptive walks in a gene network model of morphogenesis: insights into the Cambrian explosion. *Int. J. Dev. Biol.* **43**, 685–693.
214. Borenstein E, Krakauer DC. 2008 An end to endless forms: epistasis, phenotype distribution bias, and nonuniform evolution. *PLoS Comput. Biol.* **4**, e1000202. (doi:10.1371/journal.pcbi.1000202)
215. Munteanu A, Solé RV. 2008 Neutrality and robustness in evo-devo: emergence of lateral inhibition. *PLoS Comput. Biol*. **4**, e1000226. (doi:10.1371/journal.pcbi.1000226)
216. Cotterell J, Sharpe J. 2010 An atlas of gene regulatory networks reveals multiple three-gene mechanisms for interpreting morphogen gradients. *Mol. Syst. Biol*. **6**, 425. (doi:10.1038/msb.2010.74)
217. Fernández P, Solé RV. 2007 Neutral fitness landscapes in signalling networks. *J. R. Soc. Interface* **4**, 41–47. (doi:10.1098/rsif.2006.0152)
218. Catalán P, Arias CF, Cuesta JA, Manrubia S. 2017 Adaptive multiscapes: an up-to-date metaphor to visualize molecular adaptation. *Biol. Direct* **12**, 1–15. (doi:10.1186/s13062-017-0178-1)
219. Deutsch A, Dormann S. 2005 *Mathematical modeling of biological pattern formation*. Boston, MA: Springer.
220. Thomson JA. 1992 *On growth and form*. London, UK: Cambridge University Press.
221. Ball P. 1999 *The self-made tapestry: pattern formation in nature*. Oxford, UK: Oxford University Press. (doi:10.1093/oso/9780198502449.001.0001)
222. Graner F, Riveline D. 2017 'The forms of tissues, or cell-aggregates': D'Arcy Thompson's influence and its limits. *Development* **144**, 4226–4237. (doi:10.1242/dev.151233)
223. Forgacs G, Newman SA. 2005 *Biological physics of the developing embryo*. New York, NY: Cambridge University Press.
224. Stolarska MA, Kim Y, Othmer HG. 2009 Multi-scale models of cell and tissue dynamics. *Phil. Trans. R. Soc. A* **367**, 3525–3553. (doi:10.1098/rsta.2009.0095)
225. Lenne PF, Trivedi V. 2022 Sculpting tissues by phase transitions. *Nat. Commun.* **13**, 664. (doi:10.1038/s41467-022-28151-9)
226. Newman SA, Bhat R. 2009 Dynamical patterning modules: a 'pattern language' for development and evolution of multicellular form. *Int. J. Dev. Biol.* **53**, 693–705. (doi:10.1387/ijdb.072481sn)
227. Newman SA. 2012 Physico-genetic determinants in the evolution of development. *Science* **338**, 217–219. (doi:10.1126/science.1222003)
228. Graner F, Glazier JA. 1992 Simulation of biological cell sorting using a two-dimensional extended Potts model. *Phys. Rev. Lett.* **69**, 2013–2016. (doi:10.1103/PhysRevLett.69.2013)
229. Brodland GW. 2004 Computational modeling of cell sorting, tissue engulfment, and related phenomena: a review. *Appl. Mech. Rev.* **57**, 47–76. (doi:10.1115/1.1583758)
230. Steinberg MS. 1975 Adhesion-guided multicellular assembly. *J. Theor. Biol.* **55**, 431–443. (doi:10.1016/S0022-5193(75)80091-9)
231. Prusinkiewicz P, Lindenmayer A. 2012 *The algorithmic beauty of plants*. Berlin, Germany: Springer Science & Business Media.
232. Kaandorp JA. 1994 *Fractal modelling: growth and form in biology*. Berlin, Germany: Springer Science & Business Media.
233. Hoyal Cuthill JF, Conway Morris S. 2014 Fractal branching organizations of Ediacaran rangeomorph fronds reveal a lost Proterozoic body plan. *Proc. Natl Acad. Sci. USA* **111**, 13122–13126. (doi:10.1073/pnas.1408542111)
234. Wolpert L. 2015 *Principles of development*. Oxford, UK: Oxford University Press.
235. Staple DB, Farhadifar R, Röper JC, Aigouy B, Eaton S, Jülicher F. 2010 Mechanics and remodelling of cell packings in epithelia. *Eur. Phys. J. E* **33**, 117–127. (doi:10.1140/epje/i2010-10677-0)
236. Bi D, Lopez JH, Schwarz JM, Manning ML. 2015 A density-independent rigidity transition in biological tissues. *Nat. Phys.* **11**, 1074–1079. (doi:10.1038/nphys3471)
237. Jablonka E, Lamb MJ. 2006 The evolution of information in the major transitions. *J. Theor. Biol.* **239**, 236–246. (doi:10.1016/j.jtbi.2005.08.038)
238. Lyon P, Keijzer F, Arendt D, Levin M. 2021 Reframing cognition: getting down to biological basics. *Phil. Trans. R. Soc. Lond. B* **376**, 20190750. (doi:10.1098/rstb.2019.0750)
239. Wagensberg J. 2000 Complexity versus uncertainty: the question of staying alive. *Biol. Philos.* **15**, 493–508. (doi:10.1023/A:1006611022472)
240. Llinás RR. 2002 *I of the vortex: from neurons to self*. Cambridge, MA: MIT Press. (doi:10.7551/mitpress/3626.001.0001)
241. Seoane LF, Solé RV. 2018 Information theory, predictability and the emergence of complex life. *R. Soc. Open Sci.* **5**, 172221. (doi:10.1098/rsos.172221)
242. Paulin MG, Cahill-Lane J. 2021 Events in early nervous system evolution. *Top. Cogn. Sci.* **13**, 25–44. (doi:10.1111/tops.12461)
243. Sarpeshkar R. 1998 Analog versus digital: extrapolating from electronics to neurobiology. *Neural Comput.* **10**, 1601–1638. (doi:10.1162/089976698300017052)
244. Sjöström PJ, Turrigiano GG, Nelson SB. 2001 Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron* **32**, 1149–1164. (doi:10.1016/s0896-6273(01)00542-6)
245. Shagrir O. 2022 *The nature of physical computation*. New York, NY: Oxford University Press.
246. Maley CJ. 2023 Analogue computation and representation. *Br. J. Philos. Sci*. **74**, 739–769. (doi:10.1086/715031)
247. Kristan WB. 2016 Early evolution of neurons. *Curr. Biol*. **26**, R949–R954. (doi:10.1016/j.cub.2016.05.030)

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

248. Jékely G. 2011 Origin and early evolution of neural circuits for the control of ciliary locomotion. *Proc. R. Soc. B* **278**, 914–922. (doi:10.1098/rspb.2010.2027)
249. Jékely G. 2021 The chemical brain hypothesis for the origin of nervous systems. *Phil. Trans. R. Soc. B* **376**, 20190761. (doi:10.1098/rstb.2019.0761)
250. McCulloch WS, Pitts W. 1943 A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys*. **5**, 115–133. (doi:10.1007/BF02478259)
251. Cowan JD. 1990 McCulloch-Pitts and related neural nets from 1943 to 1989. *Bull. Math. Biol.* **52**, 73–97. (doi:10.1016/S0092-8240(05)80005-9)
252. Anderson JA, Rosenfeld E. 2000 *Talking nets: an oral history of neural networks*. Cambridge, MA: MIT Press.
253. Churchland PS, Sejnowski TJ. 1992 *The computational brain*. Cambridge, MA: MIT press.
254. Rojas R. 2013 *Neural networks: a systematic introduction*. Berlin, Germany: Springer Science and Business Media.
255. Fukushima K. 1980 Neocognitron: a self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybern*. **36**, 193–202. (doi:10.1007/BF00344251)
256. Glass L, Kauffman SA. 1973 The logical analysis of continuous, non-linear biochemical control networks. *J. Theor. Biol.* **39**, 103–129. (doi:10.1016/0022-5193(73)90208-7)
257. Kurten KE. 1988 Correspondence between neural threshold networks and Kauffman Boolean cellular automata. *J. Phys. A Math. Gen.* **21**, L615–L619. (doi:10.1088/0305-4470/21/11/009)
258. Luque B, Solé RV. 1997 Phase transitions in random networks: simple analytic determination of critical points. *Phys. Rev. E* **55**, 257–260. (doi:10.1103/PhysRevE.55.257)
259. Bornholdt S. 2008 Boolean network models of cellular regulation: prospects and limitations. *J. R. Soc. Interface* **5**, S85–S94. (doi:10.1098/rsif.2008.0132.focus)
260. Parisi G. 1990 A simple model for the immune network. *Proc. Natl Acad. Sci. USA* **87**, 429–433. (doi:10.1073/pnas.87.1.429)
261. Perelson AS, Weisbuch G. 1997 Immunology for physicists. *Rev. Mod. Phys.* **69**, 1219–1268. (doi:10.1103/RevModPhys.69.1219)
262. Agliari E, Barra A, Del Ferraro G, Guerra F, Tantari D. 2015 Anergy in self-directed B lymphocytes: a statistical mechanics perspective. *J. Theor. Biol.* **375**, 21–31. (doi:10.1016/j.jtbi.2014.05.006)
263. Deneubourg JL, Goss S, Franks N, Sendova-Franks A, Detrain C, Chrétien L. 1991 The dynamics of collective sorting robot-like ants and ant-like robots. In *From animals to animats*, pp. 356–365. Cambridge, MA: MIT Press. (doi:10.7551/mitpress/3115.003.0048)
264. Solé RV, Miramontes O, Goodwin BC. 1993 Oscillations and chaos in ant societies. *J. Theor. Biol.* **161**, 343–357. (doi:10.1006/jtbi.1993.1060)
265. Theraulaz G *et al*. 2002 Spatial patterns in ant colonies. *Proc. Natl Acad. Sci. USA* **99**, 9645–9649. (doi:10.1073/pnas.152302199)
266. Couzin ID. 2009 Collective cognition in animal groups. *Trends Cogn. Sci.* **13**, 36–43. (doi:10.1016/j.tics.2008.10.002)
267. McMillen P, Levin M. 2024 Collective intelligence: a unifying concept for integrating biology across scales and substrates. *Commun. Biol.* **7**, 378. (doi:10.1038/s42003-024-06037-4)
268. Yusufaly TI, Boedicker JQ. 2017 Mapping quorum sensing onto neural networks to understand collective decision making in heterogeneous microbial communities. *Phys. Biol.* **14**, 046002. (doi:10.1088/1478-3975/aa7c1e)
269. Solé R, Moses M, Forrest S. Liquid brains, solid brains. *Phil. Trans. R. Soc. B* **374**, 20190040. (doi:10.1098/rstb.2019.0040)
270. Piñero J, Solé R. 2019 Statistical physics of liquid brains. *Phil. Trans. R. Soc. B* **374**, 20180376. (doi:10.1098/rstb.2018.0376)
271. Mjolsness E, Sharp DH, Reinitz J. 1991 A connectionist model of development. *J. Theor. Biol*. **152**, 429–453. (doi:10.1016/s0022-5193(05)80391-1)
272. Alon U. 2019 *An introduction to systems biology: design principles of biological circuits*. Boca Raton, FL: CRC Press.
273. Sneppen K. 2014 *Models of life*. Cambridge, UK: Cambridge University Press.
274. Ginsburg S, Jablonka E. 2010 The evolution of associative learning: a factor in the cambrian explosion. *J. Theor. Biol*. **266**, 11–20. (doi:10.1016/j.jtbi.2010.06.017)
275. Piccinini G. 2004 The first computational theory of mind and brain: a close look at Mcculloch and Pitts's 'logical calculus of ideas immanent in nervous activity'. *Synthese* **141**, 175–215. (doi:10.1023/B:SYNT.0000043018.52445.3e)
276. Shigeno S. 2017 Brain evolution as an information flow designer: the ground architecture for biological and artificial general intelligence. In *Brain evolution by design: from neural origin to cognitive architecture*, pp. 415–438. Tokyo, Japan: Springer. (doi:10.1007/978-4-431-56469-0_19)
277. Shigeno S, Andrews PLR, Ponte G, Fiorito G. 2018 Cephalopod brains: an overview of current knowledge to facilitate comparison with vertebrates. *Front. Physiol*. **9**, 952. (doi:10.3389/fphys.2018.00952)
278. Garcia-Lopez P, Garcia-Marin V, Freire M. 2010 The histological slides and drawings of Cajal. *Front. Neuroanat*. **4**, 9. (doi:10.3389/neuro.05.009.2010)
279. North G, Greenspan RJ. 2007 *Invertebrate neurobiology*. New York, NY: Cold Spring Harbor Laboratory Press.
280. Seth AK, Bayne T. 2022 Theories of consciousness. *Nat. Rev. Neurosci*. **23**, 439–452. (doi:10.1038/s41583-022-00587-4)
281. Dennett DC. 2008 *Kinds of minds: toward an understanding of consciousness*. New York, NY: Basic Books.
282. Powell R, Mikhalevich I, Logan C, Clayton NS. 2017 Convergent minds: the evolution of cognitive complexity in nature. *Interface Focus* **7**, 20170029. (doi:10.1098/rsfs.2017.0029)
283. Dennett DC. 2017 *From bacteria to Bach and back: the evolution of minds*. New York, NY: WW Norton & Company.
284. Solé R, Seoane LF. 2022 Evolution of brains and computers: the roads not taken. *Entropy* **24**, 665. (doi:10.3390/e24050665)
285. Margalef R. 1997 *Our biosphere*. Oakland, CA: University of California Press.
286. Ghilarov AM. 2000 Ecosystem functioning and intrinsic value of biodiversity. *Oikos* **90**, 408–412. (doi:10.1034/j.1600-0706.2000.900222.x)
287. Odum EP *et al*. 1971 *Fundamentals of ecology*. Philadelphia, PA: Saunders.
288. May RM. 1973 Qualitative stability in model ecosystems. *Ecology* **54**, 638–641. (doi:10.2307/1935352)
289. Marzloff MP, Dambacher JM, Johnson CR, Little LR, Frusher SD. 2011 Exploring alternative states in ecological systems with a qualitative analysis of community feedback. *Ecol. Modell.* **222**, 2651–2662. (doi:10.1016/j.ecolmodel.2011.03.040)
290. Poulin R. 2011 *Evolutionary ecology of parasites*. Princeton, NJ: Princeton University Press.
291. Sasselov DD, Grotzinger JP, Sutherland JD. 2020 The origin of life as a planetary phenomenon. *Sci. Adv.* **6**, eaax3419. (doi:10.1126/sciadv.aax3419)
292. Batalha NM. 2014 Exploring exoplanet populations with NASA's Kepler Mission. *Proc. Natl Acad. Sci. USA* **111**, 12647–12654. (doi:10.1073/pnas.1304196111)
293. McKay CP. 2014 Requirements and limits for life in the context of exoplanets. *Proc. Natl Acad. Sci. USA* **111**, 12628–12633. (doi:10.1073/pnas.1304212111)
294. Harte J, Newman EA. 2014 Maximum information entropy: a foundation for ecological theory. *Trends Ecol. Evol.* **29**, 384–389. (doi:10.1016/j.tree.2014.04.009)
295. Gause GF. 1932 Experimental studies on the struggle for existence. *J. Exp. Biol*. **9**, 389–402. (doi:10.1242/jeb.9.4.389)
296. May RM. 1972 Will a large complex system be stable? *Nature* **238**, 413–414. (doi:10.1038/238413a0)
297. Pimm SL. 1991 *The balance of nature? Ecological issues in the conservation of species and communities*. Chicago, IL: University of Chicago Press.
298. Roopnarine PD. 2006 Extinction cascades and catastrophe in ancient food webs. *Paleobiology* **32**, 1–19. (doi:10.1666/05008.1)
299. Mitchell JS, Roopnarine PD, Angielczyk KD. 2012 Late cretaceous restructuring of terrestrial communities facilitated the end-cretaceous mass extinction in North America. *Proc. Natl Acad. Sci. USA* **109**, 18857–18861. (doi:10.1073/pnas.1202196109)
300. Adami C. 1998 *Introduction to artificial life*. New York, NY: Springer Science & Business Media.

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

301. Kristian L, Nordahl MG. 1994 Artificial food webs. In *Artificial life III*. Santa Fe, NM: Addison-Wesley.
302. Banzhaf W. 1994 Self-organisation in a system of binary strings. In *Artificial life IV proceedings of the fourth international workshop on the synthesis and samularaon of living systems* (eds RA Brooks, P Maes), pp. 109–118. Cambridge, MA: MIT Press. (doi:10.7551/mitpress/1428.003.0015)
303. Dunne JA, Williams RJ, Martinez ND, Wood RA, Erwin DH. 2008 Compilation and network analyses of cambrian food webs. *PLoS Biol*. **6**, e102. (doi:10.1371/journal.pbio.0060102)
304. Barricelli NA. 1962 Numerical testing of evolution theories: part I theoretical introduction and basic tests. *Acta Biotheor.* **16**, 69–98. (doi:10.1007/BF01556771)
305. Barricelli NA. 1963 Numerical testing of evolution theories: part II preliminary tests of performance. symbiogenesis and terrestrial life. *Acta Biotheor.* **16**, 99–126. (doi:10.1007/BF01556602)
306. Dyson G. 2012 *Turing's cathedral: the origins of the digital universe*. New York, NY: Pantheon Books.
307. Langton CG. 1986 Studying artificial life with cellular automata. *Physica D* **22**, 120–149. (doi:10.1016/0167-2789(86)90237-X)
308. Rocha L. 1997 *Evolutionary systems and artificial life*. Los Alamos, NM: Los Alamos National Laboratory.
309. Rasmussen S, Knudsen C, Feldberg R, Hindsholm M. 1990 The coreworld: emergence and evolution of cooperative structures in a computational chemistry. *Physica D* **42**, 111–134. (doi:10.1016/0167-2789(90)90070-6)
310. Solé R, Elena SF. 2018 *Viruses as complex adaptive systems*. Princeton, NJ: Princeton University Press. (doi:10.23943/princeton/9780691158846.001.0001)
311. Lloyd AL, May RM. 2001 How viruses spread among computers and people. *Science* **292**, 1316–1317. (doi:10.1126/science.1061076)
312. Ray TS. 1991 An approach to the synthesis of life. In *Artificial life II: Santa Fe Institute studies in the sciences of complexity*, pp. 371–408. Redwood City, CA: Addison-Wesley.
313. Ray TS. 1994 Evolution, complexity, entropy and artificial reality. *Physica D* **75**, 239–263. (doi:10.1016/0167-2789(94)90286-0)
314. Hillis WD. 1990 Co-evolving parasites improve simulated evolution as an optimization procedure. *Physica D* **42**, 228–234. (doi:10.1016/0167-2789(90)90076-2)
315. Dennett DC. 1995 *Darwin's dangerous idea*. New York, NY: Simon and Schuster.
316. Wilke CO, Adami C. 2002 The biology of digital organisms. *Trend. Ecol. Evol.* **17**, 528–532. (doi:10.1016/S0169-5347(02)02612-5)
317. Solé RV, Valverde S. 2013 Macroevolution *in silico*: scales, constraints and universals. *Palaeontology* **56**, 1327–1340. (doi:10.1111/pala.12047)
318. Eldredge N, Gould SJ. 1972 Punctuated equilibria: an alternative to phyletic gradualism. In *Models in paleobiology*, pp. 82–115. San Francisco, CA: Freeman, Cooper.
319. Lindgren K. 1991 Evolutionary phenomena in simple dynamics. In *Artificial life I*, pp. 295–312. Boca Raton, FL: Addison-Wesley.
320. Dittrich P, Ziegler J, Banzhaf W. 2001 Artificial chemistries: a review. *Artif. Life* **7**, 225–275. (doi:10.1162/106454601753238636)
321. Banzhaf W, Yamamoto L. 2015 *Artificial chemistries*. Cambridge, MA: MIT Press.
322. Stadler PF, Fontana W, Miller JH. 1993 Random catalytic reaction networks. *Physica D* **63**, 378–392. (doi:10.1016/0167-2789(93)90118-K)
323. Eigen M. 2013 *From strange simplicity to complex familiarity: a treatise on matter, information, life and thought*. Oxford, UK: Oxford University Press.
324. Smith E, Krishnamurthy S. 2015 *Symmetry and collective fluctuations in evolutionary games*. Bristol, UK: IOP Press.
325. Fontana W, Wagner G, Buss LW. 1993 Beyond digital naturalism. *Artif. Life* **1**, 211–227. (doi:10.1162/artl.1993.1.1_2.211)
326. Fontana W, Buss LW. 1996 The barrier of objects: from dynamical systems to bounded organizations. In *Boundaries and barriers* (eds J Casti, A Karlqvist), pp. 56–116. New York, NY: Addison-Wesley.
327. Eigen M. 1971 Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **58**, 465–523. (doi:10.1007/BF00623322)
328. Eigen M, Schuster P. 1977 The hypercycle, part A: the emergence of the hypercycle. *Naturwissenschaften* **64**, 541–565. (doi:10.1007/BF00450633)
329. Eigen M, Schuster P. 1978 The hypercycle, part C: the realistic hypercycle. *Naturwissenschaften* **65**, 341–369. (doi:10.1007/BF00439699)
330. Iranzo J, Puigbò P, Lobkovsky AE, Wolf YI, Koonin EV. 2016 Inevitability of genetic parasites. *Genome Biol. Evol.* **8**, 2856–2869. (doi:10.1093/gbe/evw193)
331. Koonin EV, Wolf YI, Katsnelson MI. 2017 Inevitability of the emergence and persistence of genetic parasites caused by evolutionary instability of parasite-free states. *Biol. Direct* **12**, 1–12. (doi:10.1186/s13062-017-0202-5)
332. Hickinbotham SJ, Stepney S, Hogeweg P. 2021 Nothing in evolution makes sense except in the light of parasitism: evolution of complex replication strategies. *R. Soc. Open Sci.* **8**, 210441. (doi:10.1098/rsos.210441)
333. Zaman L, Meyer JR, Devangam S, Bryson DM, Lenski RE, Ofria C. 2014 Coevolution drives the emergence of complex traits and promotes evolvability. *PLoS Biol*. **12**, e1002023. (doi:10.1371/journal.pbio.1002023)
334. Seoane LF, Solé R. 2023 How turing parasites expand the computational landscape of digital life. *Phys. Rev. E* **108**, 044407. (doi:10.1103/PhysRevE.108.044407)
335. Elena SF, Sanjuán R. 2007 Virus evolution: insights from an experimental approach. *Annu. Rev. Ecol. Evol. Syst*. **38**, 27–52. (doi:10.1146/annurev.ecolsys.38.091206.095637)
336. Elena SF. 2016 Evolutionary transitions during RNA virus experimental evolution. *Phil. Trans. R. Soc. B* **371**, 20150441. (doi:10.1098/rstb.2015.0441)
337. Forterre P. 2006 The origin of viruses and their possible roles in major evolutionary transitions. *Virus Res*. **117**, 5–16. (doi:10.1016/j.virusres.2006.01.010)
338. Koonin EV. 2016 Viruses and mobile elements as drivers of evolutionary transitions. *Phil. Trans. R. Soc. B* **371**, 20150442. (doi:10.1098/rstb.2015.0442)
339. Hudson PJ, Dobson AP, Lafferty KD. 2006 Is a healthy ecosystem one that is rich in parasites? *Trends Ecol. Evol.* **21**, 381–385. (doi:10.1016/j.tree.2006.04.007)
340. Kempes CP, Follows MJ, Smith H, Graham H, House CH, Levin SA. 2021 Generalized stoichiometry and biogeochemistry for astrobiological applications. *Bull. Math. Biol*. **83**, 73. (doi:10.1007/s11538-021-00877-5)
341. Gagler DC, Karas B, Kempes CP, Malloy J, Mierzejewski V, Goldman AD, Kim H, Walker SI. 2022 Scaling laws in enzyme function reveal a new kind of biochemical universality. *Proc. Natl Acad. Sci. USA* **119**, e2106655119. (doi:10.1073/pnas.2106655119)
342. Brown JH, Gillooly JF, Allen AP, Savage VM, West GB. 2004 Toward a metabolic theory of ecology. *Ecology* **85**, 1771–1789. (doi:10.1890/03-9000)
343. Szathmáry E, Smith JM. 1995 The major evolutionary transitions. *Nature* **374**, 227–232. (doi:10.1038/374227a0)
344. Szathmáry E. 2015 Toward major evolutionary transitions theory 2.0. *Proc. Natl Acad. Sci. USA* **112**, 10104–10111. (doi:10.1073/pnas.1421398112)
345. Wolf YI, Katsnelson MI, Koonin EV. 2018 Physical foundations of biological complexity. *Proc. Natl Acad. Sci. USA* **115**, E8678–E8687. (doi:10.1073/pnas.1807890115)
346. Eldredge N. 1985 *Unfinished synthesis: biological hierarchies and modern evolutionary thought*. NY: Oxford University Press. (doi:10.1093/oso/9780195036336.001.0001)
347. Tëmkin I, Eldredge N. 2015 Networks and hierarchies: approaching complexity in evolutionary theory. In *Macroevolution: explanation, interpretation and evidence* (eds E Serrelli, N Gontier), pp. 183–226. Heidelberg, Germany: Springer. (doi:10.1007/978-3-319-15045-1_6)
348. DeLong JP, Okie JG, Moses ME, Sibly RM, Brown JH. 2010 Shifts in metabolic scaling, production, and efficiency across major evolutionary transitions of life. *Proc. Natl Acad. Sci. USA* **107**, 12941–12945. (doi:10.1073/pnas.1007783107)
349. Kempes CP, Wang L, Amend JP, Doyle J, Hoehler T. 2016 Evolutionary tradeoffs in cellular composition across diverse bacteria. *ISME J.* **10**, 2145–2157. (doi:10.1038/ismej.2016.21)
350. Foëx G, Weiss P. 1926 *Le magnétisme*. Paris, France: Armand Colin.
351. Kochmański M, Paszkiewicz T, Wolski S. 2013 Curie–Weiss magnet—a simple model of phase transition. *Eur. J. Phys.* **34**, 1555–1573. (doi:10.1088/0143-0807/34/6/1555)
352. Landau LD. 1937 Theory of phase transformations. *Zh. Eksp. Teor. Fiz.* **7**, 19–32. (doi:10.1016/B978-0-08-010586-4.50034-1)

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024

353. Gell-Mann M, Low FE. 1954 Quantum electrodynamics at small distances. *Phys. Rev.* **95**, 1300–1312. (doi:10.1103/PhysRev.95.1300)
354. Wilson KG, Kogut J. 1974 The renormalization group and the ε expansion. *Phys. Rep. Phys. Lett.* **12C**, 75–200. (doi:10.1016/0370-1573(74)90023-4)
355. Weinberg S. 1979 Phenomenological lagrangians. *Physica A* **96**, 327–340. (doi:10.1016/0378-4371(79)90223-1)
356. Polchinski JG. 1984 Renormalization and effective lagrangians. *Nucl. Phys. B* **231**, 269–295. (doi:10.1016/0550-3213(84)90287-6)
357. Wiener N. 1965 *Cybernetics: or the control and communication in the animal and the machine*, 2nd edn. Cambridge, MA: MIT Press.
358. Haken H. 2010 *Synergetics: introduction and advanced topics*, 5th edn. NY: Springer Verlag.
359. Li R, Bowerman B. 2010 Symmetry breaking in biology. *Cold Spring Harb. Perspect. Biol*. **2**, a003475. (doi:10.1101/cshperspect.a003475)
360. Krakauer DC. 2023 Symmetry-simplicity, broken symmetry-complexity. *Interface Focus* **13**, 20220075. (doi:10.1098/rsfs.2022.0075)
361. Wilson KG. 1979 Problems in physics with many scales of length. *Sci. Am*. **241**, 158–179. (doi:10.1038/scientificamerican0879-158)
362. Yeomans JM. 1992 *Statistical mechanics of phase transitions*. Oxford, UK: Clarendon Press.
363. Hughes RI. 1999 The Ising model, computer simulation, and universal physics. *Ideas Context* **52**, 97–145.
364. Wikipedia. *Chirality (chemistry)*. Wikipedia, the free encyclopedia. See https://en.wikipedia.org/wiki/Chirality.
365. Farmer JD, Kauffman SA, Packard NH. 1986 Autocatalytic replication of polymers. *Physica D* **22**, 50–67. (doi:10.1016/0167-2789(86)90233-2)
366. Solé R. 2011 *Phase transitions*. Princeton, NJ: Princeton University Press.
367. Leuthäusser I. 1986 An exact correspondence between Eigen's evolution model and a two-dimensional Ising system. *J. Chem. Phys.* **84**, 1884–1885. (doi:10.1063/1.450436)
368. Leuthäusser I. 1987 Statistical mechanics of Eigen's evolution model. *J. Stat. Phys.* **48**, 343–360. (doi:10.1007/BF01010413)
369. Tarazona P. 1992 Error thresholds for molecular quasispecies as phase transitions: from simple landscapes to spin-glass models. *Phys. Rev. A* **45**, 6038–6050. (doi:10.1103/PhysRevA.45.6038)
370. Duke TAJ, Bray D. 1999 Heightened sensitivity of a lattice of membrane receptors. *Proc. Natl Acad. Sci. USA* **96**, 10104–10108. (doi:10.1073/pnas.96.18.10104)
371. Weber M, Buceta J. 2016 The cellular Ising model: a framework for phase transitions in multicellular environments. *J. R. Soc. Interface* **13**, 20151092. (doi:10.1098/rsif.2015.1092)
372. Simpson K, L'Homme A, Keymer J, Federici F. 2023 Spatial biology of Ising-like synthetic genetic networks. *BMC Biol*. **21**, 185. (doi:10.1186/s12915-023-01681-4)
373. Schlicht R, Iwasa Y. 2004 Forest gap dynamics and the Ising model. *J. Theor. Biol*. **230**, 65–75. (doi:10.1016/j.jtbi.2004.04.027)
374. De las Cuevas G, Cubitt TS. 2016 Simple universal models capture all classical spin physics. *Science* **351**, 1180–1183. (doi:10.1126/science.aab3326)
375. Fraiman D, Balenzuela P, Foss J, Chialvo DR. 2009 Ising-like dynamics in large-scale functional brain networks. *Phys. Rev. E* **79**, 061922. (doi:10.1103/PhysRevE.79.061922)
376. Nowak MA, Ohtsuki H. 2008 Prevolutionary dynamics and the origin of evolution. *Proc. Natl Acad. Sci. USA* **105**, 14924–14927. (doi:10.1073/pnas.0806714105)
377. Eigen M. 2000 Natural selection: a phase transition? *Biophys. Chem.* **85**, 101–123. (doi:10.1016/s0301-4622(00)00122-8)
378. Saito Y, Hyuga H. 2013 Colloquium: homochirality: symmetry breaking in systems driven far from equilibrium. *Rev. Mod. Phys.* **85**, 603–621. (doi:10.1103/RevModPhys.85.603)
379. Jafarpour F, Biancalani T, Goldenfeld N. 2015 Noise-induced mechanism for biological homochirality of early life self-replicators. *Phys. Rev. Lett*. **115**, 158101. (doi:10.1103/PhysRevLett.115.158101)
380. Ribó JM, Hochberg D, Crusats J, El-Hachemi Z, Moyano A. 2017 Spontaneous mirror symmetry breaking and origin of biological homochirality. *J. R. Soc. Interface* **14**, 20170699. (doi:10.1098/rsif.2017.0699)
381. Kauffman SA. 1986 Autocatalytic sets of proteins. *J. Theor. Biol*. **119**, 1–24. (doi:10.1016/s0022-5193(86)80047-9)
382. Kauffman S. 1993 *The origins of order: self-organization and selection in evolution*. New York, NY: Oxford University Press.
383. Hanel R, Kauffman SA, Thurner S. 2005 Phase transition in random catalytic networks. *Phys. Rev. E* **72**, 036117. (doi:10.1103/PhysRevE.72.036117)
384. Xavier JC, Hordijk W, Kauffman S, Steel M, Martin WF. 2020 Autocatalytic chemical networks at the origin of metabolism. *Proc. R. Soc. B* **287**, 20192377. (doi:10.1098/rspb.2019.2377)
385. Vetsigian K, Woese C, Goldenfeld N. 2006 Collective evolution and the genetic code. *Proc. Natl Acad. Sci. USA* **103**, 10696–10701. (doi:10.1073/pnas.0603780103)
386. Woese CR. 1967 *The genetic code: the molecular basis for genetic expression*. NY: Harper and Row.
387. Crick FHC. 1968 The origin of the genetic code. *J. Mol. Biol.* **38**, 367–379. (doi:10.1016/0022-2836(68)90392-6)
388. Crutchfield JP. 2012 Between order and chaos. *Nat. Phys.* **8**, 17–24. (doi:10.1038/nphys2190)
389. Langton CG. 1990 Computation at the edge of chaos: phase transitions and emergent computation. *Physica D* **42**, 12–37. (doi:10.1016/0167-2789(90)90064-V)
390. Tlusty T. 2007 A model for the emergence of the genetic code as a transition in a noisy information channel. *J. Theor. Biol.* **249**, 331–342. (doi:10.1016/j.jtbi.2007.07.029)
391. Schuster P. 1997 Genotypes with phenotypes: adventures in an RNA toy world. *Biophys. Chem.* **66**, 75–110. (doi:10.1016/s0301-4622(97)00058-6)
392. Kamp C, Bornholdt S. 2002 Coevolution of quasispecies: B-cell mutation rates maximize viral error catastrophes. *Phys. Rev. Lett.* **88**, 068104. (doi:10.1103/PhysRevLett.88.068104)
393. Chialvo DR. 2010 Emergent complex neural dynamics. *Nat. Phys.* **6**, 744–750. (doi:10.1038/nphys1803)
394. Plenz D, Niebur E (eds). 2014 *Criticality in neural systems*. Weinheim, Germany: John Wiley & Sons. (doi:10.1002/9783527651009)
395. Solé RV, Alonso D, McKane A. 2002 Self-organized instability in complex ecosystems. *Phil. Trans. R. Soc. Lond. B* **357**, 667–671. (doi:10.1098/rstb.2001.0992)
396. Biroli G, Bunin G, Cammarota C. 2018 Marginally stable equilibria in critical ecosystems. *New J. Phys*. **20**, 083051. (doi:10.1088/1367-2630/aada58)
397. Beggs JM. 2008 The criticality hypothesis: how local cortical networks might optimize information processing. *Phil. Trans. R. Soc. A* **366**, 329–343. (doi:10.1098/rsta.2007.2092)
398. Shew WL, Plenz D. 2013 The functional benefits of criticality in the cortex. *Neuroscientist* **19**, 88–100. (doi:10.1177/1073858412445487)
399. Mora T, Bialek W. 2011 Are biological systems poised at criticality? *J. Stat. Phys.* **144**, 268–302. (doi:10.1007/s10955-011-0229-4)
400. Touchette H. 2009 The large deviation approach to statistical mechanics. *Phys. Rep.* **478**, 1–69. (doi:10.1016/j.physrep.2009.05.002)
401. Smith E. 2013 Emergent order in processes: the interplay of complexity, robustness, correlation, and hierarchy in the biosphere. In *Complexity and the arrow of time* (eds C Lineweaver, P Davies, M Ruse), pp. 191–223. Cambridge, MA: Cambridge University Press. (doi:10.1017/CBO9781139225700.012)
402. Batterman RW. 2002 *The devil in the details: asymptotic reasoning in explanation, reduction, and emergence*. NY: Oxford University Press. (doi:10.1093/0195146476.001.0001)
403. Lawson-Keister E, Manning ML. 2021 Jamming and arrest of cell motion in biological tissues. *Curr. Opin. Cell Biol.* **72**, 146–155. (doi:10.1016/j.ceb.2021.07.011)
404. Hannezo E, Heisenberg CP. 2022 Rigidity transitions in development and disease. *Trends Cell Biol.* **32**, 433–444. (doi:10.1016/j.tcb.2021.12.006)
405. Corominas-Murtra B, Petridou NI. 2021 Viscoelastic networks: forming cells and tissues. *Front. Phys.* **9**, 666916. (doi:10.3389/fphy.2021.666916)
406. Petridou NI, Corominas-Murtra B, Heisenberg CP, Hannezo E. 2021 Rigidity percolation uncovers a structural basis for embryonic tissue phase transitions. *Cell* **184**, 1914–1928. (doi:10.1016/j.cell.2021.02.017)
407. Kim JH *et al*. 2020 Unjamming and collective migration in MCF10A breast cancer cell lines. *Biochem. Biophys. Res. Commun.* **521**, 706–715. (doi:10.1016/j.bbrc.2019.10.188)
408. Lenton TM, Kohler TA, Marquet PA, Boyle RA, Crucifix M, Wilkinson DM, Scheffer M. 2021 Survival of the systems. *Trends Ecol. Evol.* **36**, 333–344. (doi:10.1016/j.tree.2020.12.003)
409. Levin SR, Scott TW, Cooper HS, West SA. 2019 Darwin's aliens. *Int. J. Astrobiol.* **18**, 1–9. (doi:10.1017/S1473550417000362)

410. Koonin EV. 2007 The cosmological model of eternal inflation and the transition from chance to biological evolution in the history of life. *Biol. Direct* **2**, 1–21. (doi:10.1186/1745-6150-2-15)
411. Schuster P. 2019 What is special about autocatalysis? *Monatsh. Chem.* **150**, 763–775. (doi:10.1007/s00706-019-02437-z)
412. Langton CG. 1989 Artificial life. In *Proc. Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems*, pp. 1–47. New York, NY: Routledge. (doi:10.4324/9780429032769)
413. Bedau MA, McCaskill JS, Packard NH, Rasmussen S, Adami C, Green DG, Ikegami T, Kaneko K, Ray TS. 2000 Open problems in artificial life. *Artif. Life* **6**, 363–376. (doi:10.1162/106454600300103683)
414. Bedau MA. 2007 Artificial life. In *Philosophy of biology*, pp. 585–603. Amsterdam, The Netherlands: Elsevier. (doi:10.1016/B978-044451543-8/50027-7)
415. Lehman J, Clune J, Misevic D. 2018 The surprising creativity of digital evolution. In *Artificial life*, pp. 55–56. Cambridge, MA: MIT Press. (doi:10.1162/isal_a_00016)

Downloaded from https://royalsocietypublishing.org/ on 13 December 2024