Opinion

What does evolution make? Learning in living lineages and machines

Benedikt Hartl [1,2,4] and Michael Levin [1,3,4,*]

How does genomic information unfold, to give rise to self-constructing living organisms with problem-solving capacities at all levels of organization? We review recent progress that unifies work in developmental genetics and machine learning (ML) to understand mapping of genes to traits. We emphasize the deep symmetries between evolution and learning, which cast the genome as instantiating a generative model. The layer of physiological computations between genotype and phenotype provides a powerful degree of plasticity and robustness, not merely complexity and indirect mapping, which strongly impacts individual and evolutionary-scale dynamics. Ideas from ML and neuroscience now provide a versatile, quantitative formalism for understanding what evolution learns and how developmental and regenerative morphogenesis interpret the deep lessons of the past to solve new problems. This emerging understanding of the informational architecture of living material is poised to impact not only genetics and evolutionary developmental biology but also regenerative medicine and synthetic morphoengineering.

Whence the endless forms most beautiful?

Living forms present three fundamental challenges to our understanding: first, they self-assemble – performing all of the decision-making needed to construct a functional, complex body while the computational material itself is being reorganized on-the-fly. Second, they reach the correct target morphology reliably, utilizing heredity mechanisms to propagate specific patterns of form and behavior through time. Crucially, third, this process is almost never hard-wired, but instead offers immense plasticity, able to complete morphogenetic tasks despite perturbations of external environment and internal components [1]. This capacity to navigate the morphospace of possible anatomies, to produce the correct final pattern in the face of novel situations, or to create something completely different (never before seen by evolution) but nevertheless coherent and adaptively functional [2], is an example of problem-solving ability in a high-dimensional latent space. This lynchpin capacity ties together fields of evolutionary developmental biology, non-equilibrium thermodynamics, computational and information science, and the emerging field of diverse intelligence. The implications of understanding the multiscale behavior of the active matter of life during embryogenesis, regeneration, and cancer suppression range across biomedicine, bioengineering, robotics, and bio-inspired artificial intelligence (AI). Central to this set of questions is the relationship between the genetically specified hardware inside cells and the resulting physiological software that produces phenotypes acted upon by selection. Given the plasticity and context-sensitive decision-making of the all-important morphogenetic layer lying between genotype and phenotype, what are useful conceptual frameworks for understanding what genomes actually do (or encode), on evolutionary and ontogenic timescales?

A new perspective: the generative genome

Recent work integrating developmental biology and computer science has provided a new model of how genetic information is encoded and decoded during evolution and embryogenesis.

[1]Allen Discovery Center at Tufts University, Medford, MA, USA
[2]Institute for Theoretical Physics, TU Wien, Vienna, Austria
[3]Wyss Institute for Biologically Inspired Engineering at Harvard University, Boston, MA, USA
[4]Both authors contributed equally to this work.

*Correspondence:
michael.levin@tufts.edu (M. Levin).

Concepts from biology, neuroscience, and AI now offer a formalizable and timely framework that could advance both theoretical and empirical research [3–5]. Here, we provide a brief overview of related ideas and ways to extend this approach in theoretical and experimental biology.

Evolution through developmental reproduction involves encoding the features and functionality necessary for high-fidelity reconstruction of an organism out of the compact form of the genome. The genome does not encode organismal traits or developmental processes directly (it encodes the protein-level hardware necessary for cell behavior and information processing). The new ideas being developed try to go beyond 'blueprint' or 'program' metaphors and instead identify the genome as compressed latent variables – analogous to Waddington's gene landscape and biologically implemented via protein-encoding sequences and gene regulatory factors – that instantiate organismal development literally as a generative model [3]. It is hypothesized that the genome comprises compressed latent variables that are shaped or encoded by evolution and natural selection, and decoded by a generative model implemented by the cells of the developing embryo that is strikingly similar to the way information is processed in biological and engineered cognitive systems [3,6]. Development can thus be interpreted as a hierarchical generative decoding process from a single cell into a mature organism that is similar but not identical to its ancestors, a reconstruction with variational adaptations and mutations, but also dynamic and flexible interpretation of past information as suitable for new contexts [6]. As we argue later, this kind of architecture enables not only fidelity (reliability) in reaching the correct species-specific target morphology, but also creative problem-solving that maximizes adaptive saliency to new scenarios, not just hardwired replication of what was successful with prior environments and genetics. The line of thinking described herein integrates across scales of space and time, identifying this flexible, creative process as conserved across evolutionary, developmental, and behavioral contexts.

Evolutionary development as generative model

In ML terms, such an architecture is reminiscent of variational autoencoders (VAEs) [7–10], two-stage artificial neural networks (ANNs) [11–15] consisting of an encoder and a decoder part (Figure 1A). The encoder and decoder are jointly trained to compress input data into a lower-dimensional bottleneck representation, from which the original data is reconstructed via decompression. Typically, the decoding-stack of VAEs occurs hierarchically, going from abstract representations through adding modular features to detailed reconstructions of the original data, closely resembling, in turn, the developmental stack of embryogenesis. The latent space can capture emergent properties of the data, such as clusters or manifolds that represent different classes or features (modularity). Generative models, including VAEs, generate data by sampling from this latent space, variational latent variables allowing for the creation of novel examples conforming but not identical to the training data (adaptation). Thus, generative processes relying on compressed latent variables arguably leverage modularity and evolvability in developmental biology. Such work speaks to the fundamental question of the meaning of 'genetics' – what precisely is encoded in the biochemical medium of the genes and how is it read out and interpreted by the active cellular material whose hardware it specifies?

Extending the autoencoder analogy

This new framework focuses on the powerful idea of interpreting the genome as latent variables of a generative model. The assertion that DNA instantiates a generative model of the organism [3] has been actively debated [4–6,16–19]. However, authors differ as to whether their proposed generative model covers: (i) the entire developmental stack from the genome to the organism, or targets (ii) a single decoding layer between the genome and a cellular phenotype. This raises
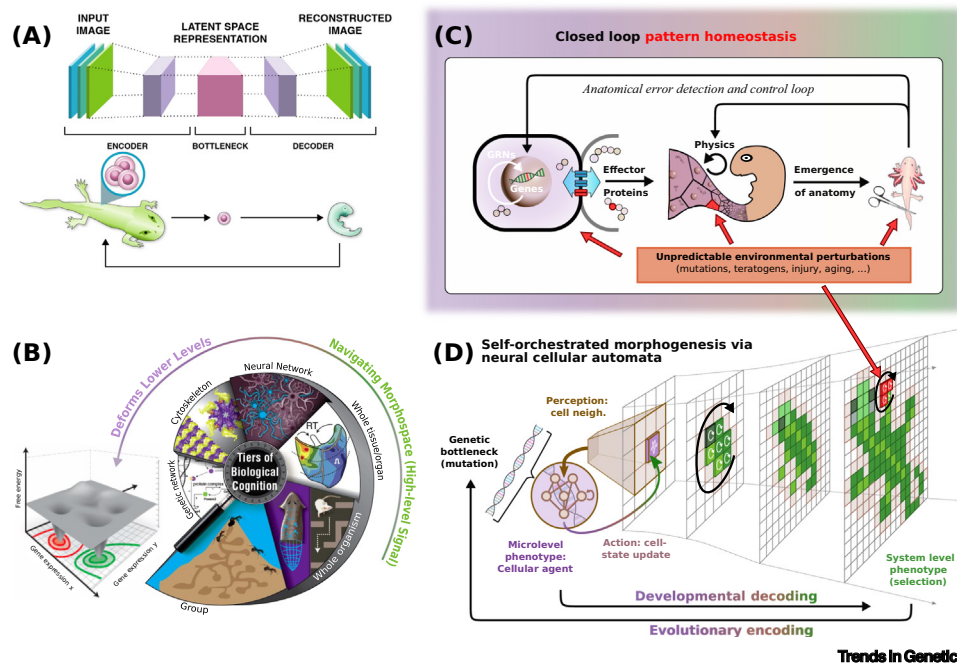
Figure 1. Developmental morphogenesis dynamically interprets genomic prompts as a multiscale generative process. (A) Evolutionary development parallels an autoencoder architecture known from machine learning, but reality is much more complicated. (B) Biology is organized as multiscale competency architecture with interactive agents navigating their respective problem spaces across scales. Lower-level components implement the physical substrate for higher levels of emergent abstraction, spanning molecular, cellular, tissue, organ, to organismal and group scales. In turn, higher levels deform the free energy landscape of possible microlevel conformations through top-down behavioral shaping and signaling. (C) Components at every level of organization perform 'local' homeostatic control loops capable of handling novel stressors on the fly, yet are susceptible to hierarchical – system level or environmental – feedback. (D) Neural cellular automata (NCA) cover the essence of our current understanding of biology's multiscale competency architecture by modeling morphogenesis through an explicitly distributed multi-agential architecture. Each cell maintains and regulates its physiological state within a local microscopic environment via successive intercellular communication and intracellular decision-making. In that way, a single cell on an NCA's grid can initiate morphogenesis of a system-level target tissue fully based on decentralized agential control loops. The NCA's cell-specific artificial neural networks models biological gene regulatory networks (GRNs) and represent the evolutionary genetic bowtie, or latent variables, of a multiscale developmental (generative) model. Panels (A), and left, central panels of (B): images by Jeremy Guay of Peregrine Creative; used with permission from [5,17,18], respectively.

questions about how such a model might be applied across different scales of biological organization (cf. Figure 1B):

(i) Biology is organized in layers within layers of abstraction, where the components of each organizational level efficiently navigate their respective problem domains [5,16,17], including metabolic, transcriptomic, physiological, anatomical, and behavioral state spaces. More than that: biological agents, even those comprising the same organism continuously influence (i.e., 'hack') each other either in symbiotic or adversarial relations, including among the organs of a single organism [20–22].

(ii) Even individual cells dynamically respond to environmental cues but can also reconfigure internally via gene regulatory networks (GRNs), displaying significant structural and operational plasticity, including several different kinds of learning [4,5,18,23].

VAEs are single-shot generators with limited generative variability stemming from noise applied at the latent space level. The subsequent decoder is a deterministic downstream process with no

room for further variability or creativity. By contrast, organismal development is inherently distributed with a collective of agents constantly reinterpreting and reacting to signals and noise across scales [5,18]. Following the VAE metaphor, this is fundamentally different from variations only in the latent space, and would rather affect every node within the decoder, massively departing from the typical deterministic downstream decoding processes. Moreover, embryonic viability requires adherence to stringent physiological and energetic criteria throughout development, while vanilla VAEs do not enforce physical constraints on intermediate states during decoding.

The hierarchical and recurrent complexity and multiplicity of the genetically instantiated generative model of organismal development poses significant challenges for current ML architectures, greatly departing from the VAE analogy. Here, we intend to expand on the more general generative model framework and discuss inherently collective self-regulatory ML approaches bridging the gap between artificial and biological life.

From bowties to Darwin's agential material

One exciting aspect of using deep concepts from ML to shed light on biology is that it provides a rigorous formalism for unifying dynamics of cognitive and morphological change. The idea that the self-assembly of the body had important symmetries with the construction of minds must have already been apparent to Turing, who famously was not only interested in unconventional embodiments of intelligence and patterns of mental activity [24] but also in the self-assembly of structural patterns during development [25]. Another notable contribution to this field was Grossberg's prescient 'Communication, memory, and development' [26]. However, advances in connectionist approaches to problem-solving in synthetic media have provided more specific architectural principles around which somatic and cognitive competency can be organized. [27–32]

Starting in 2016, Pezzulo and Levin [4] discussed how top-down approaches, such as the free energy principle (FEP) [33–36], can model developmental processes through nested bowtie architectures (i.e., where diverse molecular mechanisms converge on critical intermediates to achieve high-level goals through distributed self-orchestration). This concept highlights the efficiency and robustness of the generative developmental and evolutionary process of biological systems. Subsequent works [16,17] further explore scale-free and hierarchical dynamics in development, emphasizing the FEP's unifying role in evolution and development and introducing the concept of biology as a multiscale competency architecture (MCA), with causal top-down and bottom-up control pathways. They specifically proposed that tissue, organs, organ systems, and other levels of biological organization arise as the generalization of compressed information at molecular levels, as occurs in multilayer ANNs, and framed regeneration as a kind of pattern completion task [37] (see Box 1 for explanations of computer science terms).

This led to the idea that each individual within an MCA is an autonomous agent, actively navigating their respective environments and problem spaces [17]. In turn, a unified approach to understanding biological self-organization is arguably rooted in localized homeostatic agents following cooperatively coordinated error correction principles (Figure 1C).

Although somatic cells have been domesticated by multicellularity, and their behavior is regulated by many exogenous factors *in vivo* to align them toward large-scale setpoints at the organ and tissue level, they navigate physiological, transcriptional, metabolic, and anatomical problem spaces under the influence of their own priors as well as of cues from the local and remote internal environment [18]. Thus, a critical missing piece in our understanding of development and evolution is the fact evolution operates on an agential material, not passive, or even merely active, matter: cells can integrate multiple modalities of sensory data and make context-sensitive decisions

Box 1. Terminology

Bowtie architecture

The bowtie architecture illustrates the cyclical encoding and decoding of information at multiple scales in biological systems [132,133]. On the left side, lineage information is compressed into DNA and instances of experience are encoded into memory engrams (on the evolutionary and individual cognitive levels, respectively). The left side represents the algorithmic generalization and compression of information. The middle node represents the 'now' moment, for embryos and cognitive beings who do not have access to the past – what they have access to is the messages left for them by prior versions of themselves. The active, dynamic interpretation of these messages to drive morphogenesis or behavior is the right side of the bowtie, representing the re-inflation of the spare memory medium into the anatomy or behavior of the individual. The right side is an example of creative problem-solving because the individual has no access or allegiance to how this information was interpreted in the past – the constriction point of generative development or learning requires evolution to build systems that use that information in the most adaptive way possible at each time, constructing outcomes on-the-fly that may or may not closely resemble past form and behavior. Living systems thus operate in a continuous loop of self-construction which has allegiance to the saliency and adaptive value, not fidelity, of the meaning of the information passed down to them; this architecture implements the incredible robustness, resilience, and plasticity of life and is itself the subject of evolutionary improvements – a feedback loop and thus an intelligence ratchet that is universal throughout the agential material of life (not limited to recent inventions such as brains).

Generative models

Generative models are a class of ML systems designed to capture the underlying probability distribution of complex data (e.g., images, text, biological structures). Unlike discriminative models that classify or predict outcomes, generative models explicitly model how the data are structured. Through that they learn how to produce or generate novel samples that resemble (rather than simply duplicate) training examples by inferring and combining (i.e., generalizing from) regularities within the data. In biological terms, morphogenesis can be seen as a 'generative model' for organismal development instantiated by a genetically encoded latent space representation of the organism (i.e., the genome).

Latent space

A latent space is a lower-dimensional mathematical space in which high-dimensional data (e.g., images, proteins, phenotypes, etc.) can be represented such that they can be restored by a corresponding decoding algorithm. Points in this space represent compressed versions of the original data and are termed 'latent variables'. Depending on the compression ratio and quality of the decoder algorithm, the original high-dimensional data can be restored with high fidelity but in general not without information loss. Thus, good latent variables often capture salient patterns or important 'features' of the data, which can be interpreted modularly by the decoder. Non-comprehensive examples of such latent spaces in statistical ML are leading principal components or the bottleneck layer of autoencoders (AEs). In biological terms, the genome (or any other informational bowtie in biological organization, such as maternal proteins, cell membrane structures, and cytoskeletal organization [134–137]) represents latent variables for organismal development.

Autoencoder

AEs are ML architectures that combine two deep neural networks, an encoder (which compresses data to the latent space) and a decoder (which reconstructs data from the latent variables). AEs approximate the true data distribution by minimizing reconstruction error through variations of the encoder and decoder parameters. Depending on the quality of the latent space encoding, AEs can be used as a generative model through variations in the latent variables, leading to decoded samples that are similar but not identical to the training data. Yet, it is critical to mention that trained AEs encode and decode data monolithically, that is, by passing the data through the network exactly once: the decoder operates without feedback or context-dependent adjustments once the latent vector is fixed. Biological systems, by contrast, involve recursive processes where local interactions (e.g., cell signaling) continuously refine outcomes over time. Current AEs cannot capture this dynamic due to their static framework.

Active inference and Markov blankets

Active inference is a Bayesian framework wherein agents (e.g., organisms or cells) minimize uncertainty by reducing surprise (i.e., the deviation between expected and observed sensory data), through actions or belief updates. This process can be formalized via Markov blankets, statistical boundaries separating internal states (beliefs/actions) from environmental causes, mediated via sensory inputs and motor outputs. In general, active inference agents aim to minimize free energy, a proxy for surprise, by aligning their expectations with observations. In biology, this mirrors developmental processes: cells act as agents that adjust, for instance, gene expressions (motor commands) based on local cues such as signaling gradients, maintaining phenotypic outcomes with viability constraints over time.

Reinforcement learning

RL is a very successful ML paradigm where an agent learns to interact with an environment by taking actions, observing subsequent states, and receiving scalar reward signals that indicate the desirability of outcomes. The agent's goal is to maximize cumulative reward over time by discovering a policy (a strategy mapping states to actions), which can be

approximated by ANNs. Although this is generally not a simple task, training such an agent boils down to finding appropriate ANN parameters that allow the agent to perform well in its environment, maximizing the cumulative reward. RL principles can operate on multiple timescales, optimizing decision-making at an individual level (within a single agent's lifetime) or even evolutionary scales (e.g., simulating adaptive strategies across generations via population-based algorithms).

Criticality, and the critical brain hypothesis

The idea of criticality stems from the physics of phase transitions: a system that is disordered at high temperatures can be highly ordered at low temperatures. The critical point marks the temperature where a disordered state becomes ordered (a famous example for this is the Ising model for magnetism, which exhibits a phase transition from a high-temperature disordered paramagnetic state into an low-temperature ordered magnetized state). Criticality describes a system poised at the boundary between order and chaos, where microscopic interactions amplify into collective, scale-invariant behavior. At this critical point, the system is exceptionally sensitive to perturbations, yet maintains robust global organization, and thus gives rise to complexity. The critical brain hypothesis [85–87] argues that biological networks, particularly the brain, naturally operate near a critical state to optimize information processing, learning, and resilience. This suggests that at criticality, brains balance stability in maintaining functional structure (e.g., memory, sensory maps) while maximizing adaptability (i.e., rapid reconfiguration in response to novel stimuli or internal dynamics). Key evidence comes from neuronal avalanches, spontaneous cascades of neural activity observed across species. These avalanches follow scale-free power-law distributions, indicating the brain self-tunes to a critical regime; arguably, this state allows maximal computational efficacy.

Neural cellular automata

NCAs are systems of cellular agents distributed on a grid, each maintaining a cell-specific numerical state vector that is updated over time. In contrast to cellular automata, with predefined local update rules, NCAs utilize ANNs to mimic the internal decision-making of biological cells. By tuning the ANN parameters – all NCA cells typically sharing the same ANN – NCAs can learn the inverse problem of self-orchestrated target pattern formation based on iterative interactions: sensory perceptions of neighboring cell states are internally processed by the cell's ANN, leading to local, cell-specific state updates that coordinate globally into a developmental process. Analogous to biological morphogenesis, this process closely resembles morphogenesis via self-regulatory diversification of cells integrated in (and getting feedback from) a higher-level collective. NCAs thus provide a better analogy for biology's decentralized agency than VAEs, which lack distributed decision-making and adaptive feedback loops.

Diffusion models

DMs generate novel data by iteratively refining noisy samples toward structured outputs through progressive denoising steps. In general, DMs are realized in two phases: first, a (diffusive) forward process progressively corrupts data by iteratively blending data with noise. Second, the DM learns to predict this noise to refine initially noisy samples through iterative denoising during the (generative) reverse process. Unlike rigid decoders, as in AEs, DMs operate as self-regulatory distributed systems where each step incorporates feedback to align potentially corrupted (high-entropy) samples with a complex (low-entropy) target distribution. Thus, DMs learn the process of generating data by refining/navigating samples in the parameter space against entropic corruption and conditional to external cues, allowing the model to capture complex, modular data distributions with remarkable fidelity. This parallels biological systems' ability to robustly generate organismal form while adapting to perturbations, a capability beyond deterministic models but intrinsic to life's 'creative bowtie' architecture that depends on finding adaptive interpretations of biophysical and genetic memories.

Attention mechanism

Attention mechanisms – in ML terminology – allow models to focus on relevant parts of data when processing inputs or generating outputs [58,138]. In other words, attention mechanisms weigh the importance of different features in data dynamically (sensitive to the given context of the input), allowing models to focus on relevant parts in the data for their downstream tasks. Through this, large language models (LLMs) can set words (or better 'tokens' – abstract, computer-interpretable numerical representations of linguistic units) in context to all other tokens of a given sequence to, for example, generate text autoregressively. The working horse behind these LLMs, the transformer architecture [58], heavily relies on the attention mechanism, which can arguably be related to modern Hopfield networks [57] and thus associative memories (a subject of ongoing research). Attention can also be applied in self-regulatory scenarios, particularly in DMs, augmenting the latter with associative memory capabilities. Attention enables context-aware processing, aligning with the distributed and conditional nature of biological development, in which cells and cell groups attend to specific kinds of signals and ignore others in a context-sensitive manner.

Hopfield networks and associative memory

Hopfield networks are energy-based neural network models that store patterns as dynamical attractor states in their energy landscape. When given a degraded or partial input, they reconstruct the closest stored pattern through iterative updates, which thus represents a form of associative memory for such patterns – a mechanism that has been used to analyze pattern completion processes such as regulative embryogenesis and regeneration [37].

The generative, self-regulatory genome

While traditional generative models like AEs provide useful analogies for biological processes (especially for the evolutionary encoding of the genome as a latent variable), the current state of this architecture fails to capture essential features of life's adaptability: non-deterministic decoding, distributed agency, recursive feedback, and context-dependent inference. NCAs, DMs, and hybrid architectures inspired by active inference or attention mechanisms better align with the nested 'creative bowtie' framework of biology, offering pathways for modeling development as an emergent, generative multiscale process. This metaphor positions NCAs, and similarly DMs, as a compelling computational analogy to the genome's role as a functional generative model in development, bridging ML and systems biology by framing gene regulation not as fixed programs, but as iterative, self-regulatory, and adaptive processes that depart structured life from noise based on contextual feedback across scales.

[38–41], and even the molecular networks inside of cells can implement several different kinds of learning, including Pavlovian conditioning [23,42,43].

The various competencies of cells and subcellular components, in terms of problem-solving, memory, and flexible, context-sensitive navigation of physiological and anatomical landscapes have been reviewed elsewhere [44–47]. We argue that an organism's genotype and phenotype are connected by an agential layer of irreducible physiological computation, fundamentally implemented by self-orchestrated biological agents that are – at all scales – capable of dealing with novel situations on the fly [5]. Moreover, such an MCA demonstratively [48–50] impacts the process of an underlying evolutionary process operating on the structural and functional plasticity of an agential substrate, leading to more efficient and robust evolutionary and developmental search dynamics, adaptability, transferability, and evolvability.

Generative bowties everywhere

Mitchell and Cheney [3] identify the genome as latent variables of a generative model for organismal development that is facilitated through GRNs and shaped by evolution. Their proposed VAE architecture leverages developmental plasticity through complexity of the connectionist decoder stack, and evolvability through variability in latent variables. By contrast and complement, our framework [4–6,16–19] emphasizes the fundamental role of goal-directed problem-solving and collective intelligence of biological systems across scales in developmental and evolutionary pathways.

The paradox of change is a profound invariant across spatio-temporal scales, materials, and problem spaces, equally affecting embryos, behaving individuals, and evolutionary lineages. If a system stays the same, it will not survive when circumstances call for novel form or function. But if it changes and adapts, it is no longer the same system and the original has likewise, to an extent, disappeared. Thus, the question of what it is that persists through evolutionary change, morphogenesis/metamorphosis, and learning can be better visualized as a process of contextual sense-making, not an object. We have visualized it as a bowtie dynamic, which is traversed by active agents continuously (Figure 1). In the case of learning, the center of the bowtie is the 'now' moment. The left side of the bowtie is the deductive, algorithmic process of learning and generalization, which their cognitive system compressed from individual instances of experience. But beings never have access to the actual past – what they have access to is the engrams stored in their brain (and possibly body [51]). Thus, the right side of the bowtie is a creative process in which the cognitive system must reconstruct the meaning of these engrams, but optimizing for saliency in the current situation, not constancy of the meaning these biophysical messages had for their past self.

The same dynamic occurs in embryonic development and evolution. Counter to the familiar model in which the genetics reproduces a map of a fixed body structure, morphogenesis is a creative, problem-solving process which utilizes information in the genome but is not a hardwired, mechanical result of it. The experience of past members of a lineage is stored in a generative seed – the

genome, which does not directly code for phenotypic features. The process of producing a functional body given the prompt of the genome is as much a sense-making effort, executed by the biomechanical, biochemical, and bioelectrical networks of protoplasm to navigate physiological and anatomical spaces as is the memory-interpretation task of neural networks in brains as an animal navigates the 3D space of behavior. We propose that what evolution builds are not fixed solutions to fixed environments, but problem-solving agents, exerting well-demonstrated competencies to use genetically provided affordances in new ways as needed (discussed in detail in [5,6]).

We have proposed that biology strongly leans in to the fact that it is operating on an unreliable substrate (guaranteed to change due to mutation over long timescales and noise on short timescales) and favors architectures that take the lessons of the past seriously, but not literally, and exhibit considerable regulative and creative ways to navigate their problem spaces. Viewing biology as operating over an active, computational, decision-making material (as opposed to fixed mappings between genetic information and phenotypic outcomes) has many implications for how we relate to it in biomedical settings (recent emphasis being on prompts and behavior-shaping communication [52]), but also for evolution. Recent computational modeling efforts showed that the competency of the material hides information from selection, leading much of the work to be done on the competency mechanisms themselves [48,49]. This results in augmented ability of the material to creatively adapt to changes in its genetics, thus accelerating the evolutionary process, and kickstarting a positive feedback loop which scales primitive competencies of living matter into the more obvious intelligence of brainy animals.

It is essential to develop a rigorous understanding of the substrate- and scale-invariant learning and creative problem-solving dynamics that tie together morphogenesis, evolution, and learning [28,29,53,54].

While the VAE architecture in particular might not yet capture such recurrent multiscale dynamics, fundamental questions remain about the differences and similarities between active inference systems, collective goal-directed agents, and (self-regulatory) generative models, especially in distributed and hierarchical systems. One might interpret every biological agent as a generative model – with an informational bowtie at its core – sampling or generating actions from a policy subordinate to a higher-level goal state. Consequently, organismal development would be characterized as a highly noisy, recurrent, and hierarchically multiplexed generative process.

Parallels between construction of body and mind

> What we have lived is nothing, what we live is a point, what we have yet to live is not yet a point, but it can be a point, which together will be and have been.
>
> [– Giordano Bruno]

The idea of seeing development as a distributed agential process can be expanded to memory and cognition [6]: a developing 'Self', capable of learning through subjective cognitive experience, is ultimately in charge of reinterpreting its own thoughts, constantly reconstructing its story without having objective ground truth of the semantics of its memories [55]. By interpreting memory as compressed embedding, remembering is an active decoding process of information captured in 'creative' bowties [6]. The process of remembering parallels a context-sensitive generative process akin to associative-memory-based attention mechanisms [56–58] or, in turn, development, rendering 'memories' to be potentially hierarchical and contextually agential, constantly reinterpreted by the same 'Self'.

The concept of an MCA is applicable across biological systems, including development, evolution, and cognitive processes: it's agents all the way up and down. In turn, this suggests a unified approach to understanding biological organization – from cells to mind – can not be explained via complexity and (nested) indirect encodings alone, but rather fundamentally rooted in collective intelligence [19].

Neural cellular automata for self-orchestrated morphogenesis

Dating back to von Neumann's self-replicating machines [59], cellular automata (CAs) have provided a foundational framework for modeling distributed biological phenomena and *Artificial Life*, such as replication, growth, and morphogenesis, simply through local interactions of neighboring cells on a discrete grid. Exemplified by, for example, Conway's *Game of Life* [60] and Wolfram's *New Kind of Science* [61], even basic update rules amongst a CA's cells can give rise to emergent system-level behavior recapitulating some aspects of the dynamic, constructive quality of life. Since explicitly programming CAs for target applications proved difficult, a relatively recent but promising extension termed neural cellular automata (NCAs) [62,63] replaces the hard-wired update rules in CAs with more flexible and generalizable ANNs. Like the bowtie continuous sense-making of life mentioned earlier, an NCA's cells make dynamic decisions about how to interpret the current state of their neighborhood in order to generate their next state. Such NCAs have been successfully trained via differentiable [63–69] and evolutionary learning frameworks [49,70] to perform, amongst others, the inverse task of targeted pattern formation (*cf.* Figure 1D).

Just like CAs, the artificial cells (A-cells) of NCAs maintain numerical cell states, analogous to differentiation of biological cells (B-cells). The A-cell's internal ANN computes how to update the cell's state based on the states of its neighbors, modeling decision-making competencies of B-cells. Mirroring the structural and functional plasticity of biological development, a single A-cell can instantiate the growth of a cybernetic tissue via recurrent and fully decentralized intercellular communication and internal computation. This leads to the formation of diverse tissue types, composite organs, etc., by this multicellular substrate until the system-level goal-morphology is achieved. Thus, the functionality of such A-cells can be evolved to "grow" diverse high-fidelity target patterns *in-silico*, operating as a generative morphogenetic model.

NCAs describe an iterative multiscale generative process capable of solving the inverse pattern formation task of embryogenesis fundamentally based on distributed, multi-agential communication protocols and local iterative error-correction principles [16]. This not only represents a generative morphogenetic process, but also parallels the dynamical self-regulatory architecture of GRNs: an A-cell's state broadly captures the physiological state of B-cells, such as gene expressions or transcription factors, while its ANN substitutes the GRN. The operational point of both networks, ANNs and GRNs, is susceptible to noise, environmental cues, cell state expressions, and memory. In this context, the NCA's ANN can be interpreted as a generative genome representing the latent variables of an evolvable multicellular substrate – an informational bowtie compressed into a single A-cell – while cell states may carry epigenetic information.

NCAs have been successfully trained to perform robust embryogenesis [49,63,64,71,72], symbiotic or parasitic behavior [66,67], and even intricate self-orchestrated computational tasks such as collective classification [69], medical imaging [73,74], or pathfinding in a maze [75]. Moreover, the cellular competencies of morphogenetic NCAs have significant implications for an underlying evolutionary process [49], facilitating enhanced evolvability and transferability to novel problems, and potentially enabling open-ended evolution through concepts from the paradox of robustness [76,77]. Moreover, the A-cells' ANNs may exhibit graph-like recurrent architectures [49], and thus potentially model active inference agents with an internal world-model, whereas cell states

represent communication boundaries or Markov blankets. Thus, NCAs are promising models for understanding the MCA of biology and exploiting it in novel ways for top-down control in biomedical [78] and bioengineering [79,80] settings.

A model for distributed decision-making

Current research aims to explicitly extend NCAs towards hierarchical architectures [81–83]: departing from locally communicating cellular neighbors on a single NCA layer, hierarchical NCAs capture emerging inter-layer structures (such as clusters of cells) as effective neighbors in successively higher layers of abstraction, thereby allowing intermodular coordination across scales. While potentially improving NCAs' applicability, this might not be necessary conceptually to achieve multiscale dynamics: at the critical point, the Ising model – the simplest CA-like model for a magnetic material – displays self-similar organization patterns of magnetic islands. An NCA operating near such a critical point may exhibit scale-free spatio-temporal activation patterns in its distribution of cell states [84], potentially leading to collective interactions of such higher-level structures. Mirroring the near-critical activation patterns observed in neuronal tissue [85–87], this yet again blurs the line between morphogenesis and behavioral information processing, which is essential for capturing the nature of living things that process information while actively remodeling the processing medium itself (see Box 1 for explanations of physics-related and computer science terms, also for the following).

While indirect encoding paradigms in the domain of reinforcement learning (RL) [88] demonstrate impressive performance [89–93], encoding innate behavior in a genomic bottleneck [94], NCAs have been utilized to facilitate both morphogenetic development and robust, decentralized, and transferable behavioral control in composite RL agents [71,95–97]: purely based on intercellular communication and intracellular information processing, NCAs can grow functional robotic morphologies (by self-assembling synthetic sensory, messenger, and motor cells) that display behavioral problem-solving capabilities at the system-level based on decentralized coordination of their components. Similarly, neural developmental programs have been used to grow an agent's controller ANN during its lifetime [98,99]. Other promising RL approaches [100–102] maintain distributed sensors whose individual recurrent perceptions are analyzed for importance by attention-based mechanisms to retain a compressed environmental representation for subsequent decision-making. Such informational bowties, be it in the single-agent or collective intelligence domain [103], or even in world-model architectures [104–106], allow correspondingly trained – or evolved – agents to generalize well to environmental rules and navigate their respective problem spaces robustly.

Diffusion models: development and evolution as incremental denoising

Although progress has been made [70], NCAs are still limited in their capacity of generating multiple patterns from different seed cells. However, NCAs are closely related to probabilistic diffusion models (DMs) [107–111], cutting-edge generative techniques that demonstrate versatility and control over – so far – mainly multimedia content generation.

Both methods iteratively refine an originally incomplete or noisy data sample through incremental error correction until it conforms to a target data distribution. Different training and denoising paradigms based on non-equilibrium thermodynamics [111,112] allow DMs to generate realistic samples that capture versatile data distributions, a process that can even be guided by text prompts [113,114]. To name but a few examples, DMs have revolutionized guided generation of synthetic images [113], videos [115], and protein-folding predictions [116], and even facilitate fully generative game-play [117] demonstrating agential behavior.

Notably, we have recently demonstrated [118,119] that DMs can serve as highly efficient generative models in evolutionary algorithms, substituting conditional genotypic offspring generation
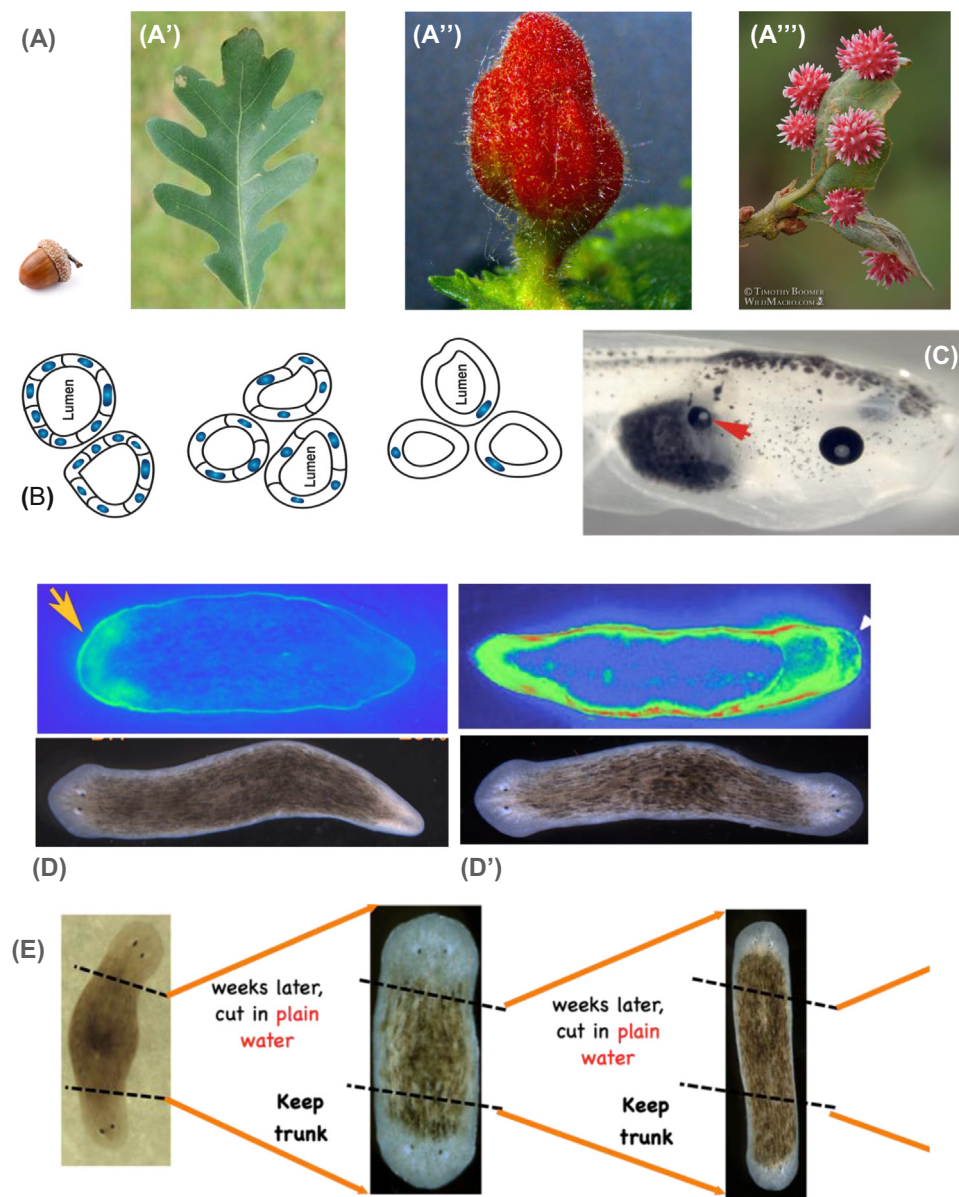
Figure 2. Plasticity of the interpretation layer between genotype and phenotype. An acorn (A) contains the genome of the oak tree, and results in a very consistent, stereotypical morphology of a species-specific leaf (A'). However, with prompts from a non-human bioengineer, such as a wasp, the plant cells are capable of making a round, spiky, red gall structure [139] instead of the normal flat green one, illustrating the plasticity of the outcome driven by the normal genome. Animal development also features remarkable examples of plasticity, in the context of problem-solving that involves accommodating not only novel aspects of the external environment but also changes in internal components. In newts (B), the normal arrangement of eight cells that create the kidney tubule changes when polyploidy causes the size of cells to increase: fewer cells are used to make the same size structure, as the system works to reach the same anatomical setpoint despite changes in available parts. The most remarkable thing happens when cells are made truly huge, in which case cytoskeletal bending – a different mechanism – is used as cells bend around themselves to leave a lumen in the middle [140]. This illustrates how morphogenesis is not a hardwired consequence of genetic information, but instead a problem-solving process that navigates anatomical morphospace using its genetic affordances as needed to achieve specific anatomical ends. Because of the unreliable material of life (an embryo cannot assume how much genetic information it will have, or what mutations may have occurred, or how many cells it will have [141], etc.), the genetic lessons of past lineages are not always taken literally

(Figure legend continued at the bottom of the next page.)

through iterative denoising of genomic parametrizations in various optimization domains. More-over, the generative process of DMs is capable of adapting a given data instance conditionally to some external cues or constraints, such as modularly replacing details in sceneries [120], transforming an image to a particular target style without corrupting its content [121], or adjusting an RL agent's policy dynamically to achieve targeted behavior shaping [119]. Thus, DMs (and similarly NCAs) are promising candidates to model the modular, internal and external context-sensitive reconfiguration capabilities of B-cells.

While the flow of information in the VAE model of evolutionary development is deterministic and unidirectional when traversing a hard-wired and predefined decoder stack – after possible variations in the latent space – both DMs and NCAs iteratively operate on dynamically re-fined structures. More than that, these increasingly contextualized structures are recursively re-introduced as refined inputs to successive generative steps, effectively conditioning and reconfiguring the functionality of their governing generative models in a dynamic, context-sensitive manner. Through a series of phase transitions from high entropy noise to novel structured samples, DMs organize data hierarchically [122–125]: initial features, emerging at early stages during the generative phase, are later refined via complexification. Details are iteratively added to initially rough sketches of the target pattern, while the order of feature refinement is modularly interchangeable [122]. Instead of a downstream latent space encoding and decoding, DMs capture the latent variables of a self-regulatory generative process, generalizing the dynamics of blending hierarchical (phenotypical) features contextually. This mirrors biological development much more accurately, where the output of computational processes is constantly fed back to the biological hardware, inducing architectural and topological changes to itself, and driving the morphological and functional plasticity of this process.

Given their generative fidelity and context-sensitive versatility, DMs potentially represent powerful models for morphogenesis and, in the spirit of this contribution, might be candidates for foundational generative models for DNA or GRNs with deep associative memory capacities [124,126]: the DM would literally represent the generative and functional genome [3,4], while the data sample it operates on relates to gene expressions or transcriptomic data [119].

Concluding remarks

The material of life exhibits active problem-solving capacities in physiological and anatomical spaces which dwarf our current efforts in biomedicine and bioengineering, in the same way that brains' learning capacities are still unmatched by efforts in ML and AI (Figures 2 and 3). An especially impressive example are asexual strains of planarian flatworms, which overcome

but can be used as resources in novel scenarios. For this reason, morphogenesis is reprogrammable – its plasticity is accessible to hacking via the interfaces cell groups use to orchestrate their journeys through anatomical morphospace. Such prompts can be bioelectrical, such as the ectopic eye on the gut of a tadpole (C, red arrowhead) induced by injection of a potassium channel [142] which establishes an eye-specific bioelectric state and can do so even outside the zone of tissue competency defined by the Pax6 'master' eye gene. In addition to specifying organ identity, bioelectric signals can control axial polarity: for example, in planaria, the normal one-headed form (D, bottom panel) possesses a bioelectric profile (D, top panel showing fluorescent voltage reporter dye signal) that indicates the number of heads. When rewritten using brief exposure to bioelectric-modifying drugs, the pattern memory can be changed into a two-headed form (D', top panel) resulting in a 2-headed worm (D', bottom panel) [143]. This change of target morphology is permanent (E), as cut two-headed worms regenerate two-headed offspring in perpetuity; such a massive change in the bodyplan of the organism does not require any genetic modification and is invisible to genomic sequencing, illustrating how form and function specification can be stored and processed in the physiological layers and can differ from the genomic default [130]. Images used with permission from references indicated, except as follows. (A') is a white oak leaf photographed by Chris Evans, River to River CWMA, Bugwood.org. (A''') – Timothy Boomer. (B) – Jeremy Guay drawn after [140]. (D) – from [144]. (E) – Junji Morokuma and Nestor Oviedo, Levin lab. iStock photo 187333325 by dionisvero, https://www.istockphoto.com/photo/one-acorn-gm187333325-28809816?searchscope=image%2Cfilm 2.

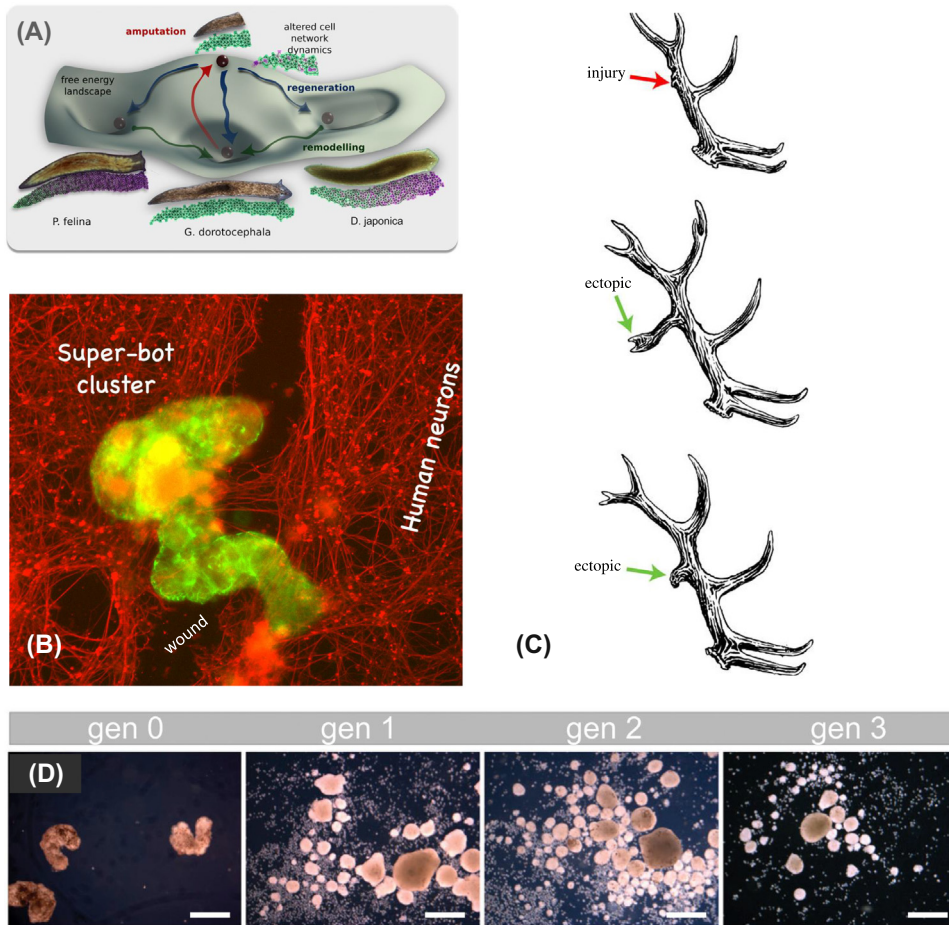Figure 3. Navigating morphospace in novel ways by reinterpreting (evolutionary) experience. The morphological plasticity in planaria flat worms (cf. Figure 2E in the main text) affects not only head number but head shape (A), with changes to the bioelectric pattern able to push genetically normal flatworm fragments to find and implement morphogenetic attractors normally occupied by other species of flatworms at 100–150 million years of evolutionary distance [128]. Anthrobots (B, showing a cluster in green) are motile proto-organisms that self-assemble from normal adult human tracheal epithelium, do not resemble any stage of human development in their form and behavior, and have the ability to heal neural wounds (B, neurons are in red; white arrow indicates Anthrobot-induced repair) [145]. In addition to being able to find coherent modes of form and function different from the species-specific default, mammalian tissues also exhibit experience-dependent plasticity of long-term anatomy. The phenomenon is known as trophic memory (C) in deer antler regeneration [146], in which injury induces ectopic tine growth in subsequent years – long after the original bone has been shed (a memory of the location of past damage that imposes dynamic changes on future morphogenesis). Similarly (D), Xenobots – motile biobots formed from frog epithelial cells [147] – exhibit novel behaviors such as kinematic self-replication. When provided with loose epithelial cells, they corral them into spheres and polish them until they too turn into Xenobots and continue the cycle (generation 1, 2, and 3 shown). Images used with permission from references indicated, except for (A) – Alexis Pietak.

states to structured data. Given the so-far limited generative storage capabilities of NCAs, an open question is whether DMs – foundational models of modern AI tools such as StableDiffusion and Midjourney with tremendous generative and associative potential – would even be better suited to understand biological genetic data from a generative perspective.

Generative models versus agents versus error-correction: it is important to understand the overlap and differences between self-regulatory generative models, multi-agential (active-inference) frameworks, and hierarchically interlocked homeostatic systems, to develop the best formalism that can exploit increasingly available multimodel large datasets from health and disease to infer intervention strategies for biomedical and bioengineering contexts.

whole body damage, cancer, and aging (despite a very noisy genome, which accumulates mutations due to somatic inheritance [127]), can implement other species' morphotypes [128] and novel, permanent bodyplan architectures in the absence of genetic change [129,130] (Figures 2D,E and 3A), and identify a small number of genes sufficient to counteract the effects of a novel toxin [131]. The challenges and opportunities facing the field (see Outstanding questions) now is to exploit and extend the deep lessons of connectionist neuroscience and ML, with

respect to architectures that learn and adapt. Exciting developments in genetics include fundamental changes in how we see the subject of the science itself, and what we are really doing when we track and edit genetic information. The molecular medicine of the future, and our understanding of natural and experimentally guided evolution, depends on developing omics tools, modeling approaches, and interventions that extend classic views of the mapping between genetics and phenotypes, and tames system-level complex outcomes by targeting the genetic, physiological, and informational structures as active learning agents.

Declaration of interests

M.L.'s lab receives funding from a sponsored research agreement with Astonishing Labs (AL), a company that seeks to operate in the space of biomedical advances driven by a better understanding of the computational intelligence and reprogrammability of living materials. M.L. also advises AL. There are currently no relevant patents but invention disclosures have been filed with Tufts University on related work on parallels between machine learning concepts and evolution (and they may or may not pursue a patent filing).

References

1. Levin, M. (2023) Collective intelligence of morphogenesis as a teleonomic process. In *Evolution "On Purpose": Teleonomy in Living Systems*, pp. 175–198, The MIT Press
2. Blackiston, D. *et al.* (2021) A cellular platform for the development of synthetic living machines. Science. *Robotics* 6, eabf1571. https://www.science.org/doi/abs/10.1126/scirobotics.abf1571
3. Mitchell, K.J. and Cheney, N. (2016) The Genomic Code: the genome instantiates a generative model of the organism. *Trends Genet.*, Published online February 10, 2025. https://dx.doi.org/10.1016/j.tig.2025.01.008
4. Pezzulo, G. and Levin, M. (2016) Top-down models in biology: explanation and control of complex living systems above the molecular level. *J. R. Soc. Interface* 13, 20160555. https://royalsocietypublishing.org/doi/10.1098/rsif.2016.0555
5. Levin, M. (2023) Darwin's agential materials: evolutionary implications of multiscale competency in developmental biology. *Cell. Mol. Life Sci.* 80, 142. https://link.springer.com/10.1007/s00018-023-04790-z
6. Levin, M. (2024) Self-Improvising Memory: A Perspective on Memories as Agential, Dynamically Reinterpreting Cognitive Glue. *Entropy* 26, 481. https://www.mdpi.com/1099-4300/26/6/481
7. Kramer, M.A. (1991) Nonlinear principal component analysis using autoassociative neural networks. *AIChE J* 37, 233–243
8. Kramer, M. (1992) Autoassociative neural networks. *Comput. Chem. Eng.* 16, 313–328
9. Kingma, D.P. and Welling, M. (2014) Auto-Encoding Variational Bayes. In *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings*
10. Higgins, I. *et al.* (2017) beta-vae: Learning basic visual concepts with a constrained variational framework. In *International conference on learning representations*
11. McCulloch, W.S. and Pitts, W. (1943) A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* 5, 115–133
12. Rosenblatt, F. (1958) The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol. Rev.* 65, 386–408
13. Rumelhart, D.E. *et al.* (1986) Learning representations by back-propagating errors. *Nature* 323, 533–536
14. LeCun, Y. *et al.* (2015) Deep learning. *Nature* 521, 436–444
15. Goodfellow, I. *et al.* (2016) *Deep Learning*, MIT Press
16. Fields, C. and Levin, M. (2020) Scale-Free Biology: Integrating Evolutionary and Developmental Thinking. *BioEssays* 42, 1900228. https://onlinelibrary.wiley.com/doi/10.1002/bies.201900228
17. Fields, C. and Levin, M. (2022) Competency in Navigating Arbitrary Spaces as an Invariant for Analyzing Cognition in Diverse Embodiments. *Entropy* 24, 819. https://www.mdpi.com/1099-4300/24/6/819
18. Levin, M. (2022) Technological Approach to Mind Everywhere: An Experimentally-Grounded Framework for Understanding Diverse Bodies and Minds. *Front. Syst. Neurosci.* 16, 768201. https://www.frontiersin.org/articles/10.3389/fnsys.2022.768201/full
19. McMillen, P. and Levin, M. (2024) Collective intelligence: A unifying concept for integrating biology across scales and substrates. *Commun. Biol.* 7, 378. https://www.nature.com/articles/s42003-024-06037-4
20. Bongard, J. and Levin, M. (2023) There's plenty of room right here: Biological systems as evolved, overloaded, multi-scale machines. *Biomimetics* 8, 110
21. Rule, J.S. *et al.* (2020) The child as hacker. *Trends Cogn. Sci.* 24, 900–915
22. Gawne, R. *et al.* (2020) Competitive and coordinative interactions between body parts produce adaptive developmental outcomes. *Bioessays* 42, e1900245
23. Biswas, S. *et al.* (2022) Learning in transcriptional network models: computational discovery of pathway-level memory and effective interventions. *Int. J. Mol. Sci.* 24, 285
24. Turing, A.M. (1950) I.—computing machinery and intelligence. *Mind* LIX, 433–460. https://doi.org/10.1093/mind/LIX.236.433
25. Turing, A.M. (1952) The chemical basis of morphogenesis. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 237, 37–72. https://royalsocietypublishing.org/doi/abs/10.1098/rstb.1952.0012
26. Grossberg, S. (1978) Communication, memory, and development. In *Progress in Theoretical Biology*, pp. 183–232, Elsevier
27. Buckley, C.L. *et al.* (2024) Natural Induction: Spontaneous Adaptive Organisation without Natural Selection. *Entropy* 26, 765. https://www.mdpi.com/1099-4300/26/9/765
28. Kouvaris, K. *et al.* (2017) How evolution learns to generalise: Using the principles of learning theory to understand the evolution of developmental organisation. *PLoS Comput. Biol.* 13, e1005358
29. Watson, R.A. *et al.* (2016) Evolutionary connectionism: Algorithmic principles underlying the evolution of biological organisation in evo-devo, evo-eco and evolutionary transitions. *Evol. Biol.* 43, 553–581

30. Watson, R.A. and Szathmáry, E. (2016) How can evolution learn? *Trends Ecol. Evol.* 31, 147–157

31. Power, D.A. *et al.* (2015) What can ecosystems learn? expanding evolutionary ecology with learning theory. *Biol. Direct* 10. https://doi.org/10.1186/s13062-015-0094-1

32. Watson, R.A. *et al.* (2014) The evolution of phenotypic correlations and "developmental memory". *Evolution* 68, 1124–1138. https://onlinelibrary.wiley.com/doi/abs/10.1111/evo.12337

33. Friston, K.J. *et al.* (2017) Active inference, curiosity and insight. *Neural Comput.* 29, 2633–2683

34. Friston, K. *et al.* (2013) The anatomy of choice: active inference and agency. *Front. Hum. Neurosci.* 7. https://doi.org/10.3389/fnhum.2013.00598

35. Friston, K. *et al.* (2012) Active inference and agency: optimal control without cost functions. *Biol. Cybern.* 106, 523–541

36. Friston, K. (2010) The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138

37. Pezzulo, G. and Levin, M. (2015) Re-membering the body: applications of computational neuroscience to the top-down control of regeneration of limbs and other complex organs. *Integr. Biol.* 7, 1487–1517

38. Zakirov, B. *et al.* (2021) Active perception during angiogenesis: filopodia speed up notch selection of tip cells *in silico* and *in vivo*. *Philos. Trans. R. Soc. B Biol. Sci.* 376, 20190753

39. Dine, E. *et al.* (2018) Protein phase separation provides long-term memory of transient spatial stimuli. *Cell Syst.* 6, 655–663.e5

40. Bugaj, L.J. *et al.* (2016) Interrogating cellular perception and decision making with optogenetic tools. *J. Cell Biol.* 216, 25–28

41. Bentley, K. *et al.* (2014) Do endothelial cells dream of eclectic shape? *Dev. Cell* 29, 146–158

42. Biswas, S. *et al.* (2021) Gene regulatory networks exhibit several kinds of memory: Quantification of memory in biological and random transcriptional networks. *iScience* 24, 102131

43. Csermely, P. *et al.* (2020) Learning of signaling networks: Molecular mechanisms. *Trends Biochem. Sci.* 45, 284–294

44. Baluška, F. *et al.* (2022) Cellular sentience as the primary source of biological order and evolution. *Biosystems* 218, 104694

45. Reber, A.S. and Baluška, F. (2021) Cognition in some surprising places. *Biochem. Biophys. Res. Commun.* 564, 150–157

46. Baluška, F. and Levin, M. (2016) On having no head: Cognition throughout biological systems. *Front. Psychol.* 7,

47. Lyon, P. (2015) The cognitive cell: bacterial behavior reconsidered. *Front. Microbiol.* 6. https://doi.org/10.3389/fmicb.2015.00264

48. Shreesha, L. and Levin, M. (2023) Cellular Competency during Development Alters Evolutionary Dynamics in an Artificial Embryogeny Model. *Entropy* 25, 131. https://www.mdpi.com/1099-4300/25/1/131

49. Hartl, B. *et al.* (2024) Evolutionary Implications of Self-Assembling Cybernetic Materials with Collective Problem-Solving Intelligence at Multiple Scales. *Entropy* 26, 532. https://www.mdpi.com/1099-4300/26/7/532

50. Caldwell, J. *et al.* (2022) Deep optimisation: Transitioning the scale of evolutionary search by inducing and searching in deep representations. SN Computer. *Science* 3. https://doi.org/10.1007/s42979-022-01109-w

51. Blackiston, D.J. *et al.* (2015) The stability of memories during brain remodeling: A perspective. *Commun. Integr. Biol.* 8, e1073424

52. Lagasse, E. and Levin, M. (2023) Future medicine: from molecular pathways to the collective intelligence of the body. *Trends Mol. Med.* 29, 687–710

53. Vanchurin, V. *et al.* (2022) Toward a theory of evolution as multilevel learning. *Proc. Natl. Acad. Sci.* 119, e2120037119. https://www.pnas.org/doi/abs/10.1073/pnas.2120037119

54. Czégel, D. *et al.* (2019) Multilevel selection as bayesian inference, major transitions in individuality as structure learning. Royal Society Open. *Science* 6, 190202. https://royalsocietypublishing.org/doi/abs/10.1098/rsos.190202

55. Kahneman, D. (2011) *Thinking, fast and slow*, Farrar, Straus and Giroux, New York ISBN 9780374275631 0374275637

56. Sacco, F. *et al.* (2023) *Requirements for self-organization*, Zenodo https://zenodo.org/record/8416764

57. Ramsauer, H. *et al.* (2021) Hopfield networks is all you need. In *International Conference on Learning Representations*

58. Vaswani, A. *et al.* (2017) Attention is all you need. In *Advances in Neural Information Processing Systems vol. 30* (Guyon, I. and Luxburg, U.V. and Bengio, S. and Wallach, H. and Fergus, R. and Vishwanathan, S. and Garnett, R., eds), Curran Associates, Inc

59. Von Neumann, J. *et al.* (1966) Theory of self-reproducing automata. *IEEE Trans. Neural Netw.* 5, 3–14

60. Games, M. (1970) The fantastic combinations of john conway's new solitaire game "life" by martin gardner. *Sci. Am.* 223, 120–123

61. Wolfram, S. (2002) *A New Kind of Science*, Wolfram Media ISBN 1579550088. https://www.wolframscience.com

62. Li, X. and Yeh, A.G.O. (2002) Neural-network-based cellular automata for simulating multiple land use changes using GIS. *Int. J. Geogr. Inf. Sci.* 16, 323–343

63. Mordvintsev, A. *et al.* (2020) Growing neural cellular automata. *Distill* 5. https://doi.org/10.23915/distill.00023

64. Randazzo, E. *et al.* (2023) *Growing steerable neural cellular automata*. vol. ALIFE 2023: Ghost in the Machine: Proceedings of the 2023 Artificial Life Conference of *Artificial Life Conference Proceedings*, p. 2. https://doi.org/10.1162/isal_a_00564

65. Mordvintsev, A. *et al.* (2022) *Growing isotropic neural cellular automata*. vol. ALIFE 2022: The 2022 Conference on Artificial Life of *Artificial Life Conference Proceedings*, p. 65. https://doi.org/10.1162/isal_a_00552

66. Cavuoti, L. *et al.* (2022) *Adversarial takeover of neural cellular automata*. vol. ALIFE 2022: The 2022 Conference on Artificial Life of *Artificial Life Conference Proceedings*, p. 38. https://doi.org/10.1162/isal_a_00521

67. Randazzo, E. *et al.* (2021) Adversarial reprogramming of neural cellular automata. *Distill* 6, http://dx.doi.org/10.23915/distill.00027.004

68. Niklasson, E. *et al.* (2021) Self-organising textures. *Distill* 6, http://dx.doi.org/10.23915/distill.00027.003

69. Randazzo, E. *et al.* (2020) Self-classifying mnist digits. *Distill* 5, http://dx.doi.org/10.23915/distill.00027.002

70. Montero, M.L. *et al.* (2024) Meta-learning an evolvable developmental encoding. In *Proceedings of the ALIFE 2024: Proceedings of the 2024 Artificial Life Conference. ALIFE 2024: Proceedings of the 2024 Artificial Life Conference*. Online. (pp. 50). ASME. https://doi.org/10.1162/isal_a_00775

71. Pontes-Filho, S. *et al.* (2022) A unified substrate for body-brain co-evolution. In *From Cells to Societies: Collective Learning across Scales*

72. Sudhakaran, S. *et al.* (2021) *Growing 3d artefacts and functional machines with neural cellular automata*. vol. ALIFE 2021: The 2021 Conference on Artificial Life of *Artificial Life Conference Proceedings*, p. 108. https://doi.org/10.1162/isal_a_00451

73. Kalkhof, J. *et al.* (2024) Unsupervised training of neural cellular automata on edge devices. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2024: 27th International Conference, Marrakesh, Morocco, October 6–10, 2024. Proceedings, Part III*, pp. 498–507, Springer-Verlag, Berlin, Heidelberg

74. Mittal, A. *et al.* (2025) Medsegdiffnca: Diffusion models with neural cellular automata for skin lesion segmentation. https://arxiv.org/abs/2501.02447

75. Earle, S. *et al.* (2023) Pathfinding neural cellular automata. https://arxiv.org/abs/2301.06820

76. Frank, S.A. (2019) Evolutionary design of regulatory control. ii. robust error-correcting feedback increases genetic and phenotypic variability. *J. Theor. Biol.* 468, 72–81

77. Frank, S.A. (2007) Maladaptation and the paradox of robustness in evolution. *PLoS One* 2, e1021

78. Levin, M. (2024) The multiscale wisdom of the body: Collective intelligence as a tractable interface for next-generation biomedicine. *BioEssays*, https://doi.org/10.1002/bies.202400196

79. Davies, J. and Levin, M. (2023) Synthetic morphology with agential materials. *Nat. Rev. Bioeng.* 1, 46–59

80. Kamm, R.D. *et al.* (2018) Perspective: The promise of multicellular engineered living systems. *APL Bioeng.* 2, 040901

81. Bielawski, K. *et al.* (2024) Evolving Hierarchical Neural Cellular Automata. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 78–86, ACM, Melbourne VIC Australia

82. Pande, R. and Grattarola, D. (2023) *Hierarchical neural cellular automata.* vol. ALIFE 2023: Ghost in the Machine: Proceedings of the 2023 Artificial Life Conference of *Artificial Life Conference Proceedings*, p. 20. https://doi.org/10.1162/isal_a_00601

83. Palm, R.B. *et al.* (2022) Variational neural cellular automata. In *International Conference on Learning Representations*

84. Manchanda, K. *et al.* (2013) Scaling behavior in probabilistic neuronal cellular automata. *Phys. Rev. E* 87, 012704. https://link.aps.org/doi/10.1103/PhysRevE.87.012704

85. Dunkelmann, S. and Radons, G. (1994) Neural networsk and abelian sandpile models of self-organized criticality. In *Proceedings of International Conference Artificial Neural Networks*, pp. 867–870, Springer-Verlag

86. Hesse, J. and Gross, T. (2014) Self-organized criticality as a fundamental property of neural systems. *Front. Syst. Neurosci.* 8. https://doi.org/10.3389/fnsys.2014.00166

87. Beggs, J.M. and Timme, N. (2012) Being critical of criticality in the brain. *Front. Physiol.* 3, https://www.frontiersin.org/journals/physiology/articles/10.3389/fphys.2012.00163

88. Sutton, R.S. and Barto, A.G. (2018) *Reinforcement Learning: An Introduction*, The MIT Press http://incompleteideas.net/book/the-book-2nd.html

89. Risi, S. *et al.* (2010) Evolving the placement and density of neurons in the hyperneat substrate. In *Proceedings of the 12th Annual Conference on Genetic and Evolutionary Computation. GECCO '10*, pp. 563–570, Association for Computing Machinery, New York, NY, USA

90. Risi, S. and Stanley, K.O. (2010) Indirectly encoding neural plasticity as a pattern of local rules. In *From Animals to Animats* (11) (Doncieux, S. and Girard, B. and Guillot, A. and Hallam, J. and Meyer, J.A. and Mouret, J.B., eds), pp. 533–543, Springer Berlin Heidelberg, Berlin, Heidelberg

91. Stanley, K.O. *et al.* (2009) A hypercube-based encoding for evolving large-scale neural networks. *Artif. Life* 15, 185–212. https://doi.org/10.1162/artl.2009.15.2.15202

92. D'Ambrosio, D.B. and Stanley, K.O. (2007) A novel generative encoding for exploiting neural network sensor and output geometry. In *Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation. GECCO '07*, pp. 974–981, Association for Computing Machinery, New York, NY, USA

93. Stanley, K.O. (2007) Compositional pattern producing networks: A novel abstraction of development. *Genet. Program Evolvable Mach.* 8, 131–162

94. Shuvaev, S. *et al.* (2024) Encoding innate ability through a genomic bottleneck. *Proc. Natl. Acad. Sci.* 121. https://doi.org/10.1073/pnas.2409160121

95. Hartl, B. *et al.* (2025) Neuroevolution of decentralized decision-making in N-bead swimmers leads to scalable and robust collective locomotion. *Commun. Phys.* 8, 194

96. Najarro, E. *et al.* (2022) HyperNCA: Growing developmental networks with neural cellular automata. In *From Cells to Societies: Collective Learning across Scales*

97. Pathak, D. *et al.* (2019) Learning to control self-assembling morphologies: a study of generalization via modularity. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, Curran Associates Inc., Red Hook, NY, USA

98. Nisioti, E. *et al.* (2024) Growing artificial neural networks for control: the role of neuronal diversity. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion (GECCO)*. Assoc. Comput. Mach. pp. 175-178. https://dl.acm.org/doi/10.1145/3638530.3654356

99. Najarro, E. *et al.* (2023) *Towards self-assembling artificial neural networks through neural developmental programs*. vol. ALIFE 2023: Ghost in the Machine: Proceedings of the 2023 Artificial Life Conference of *Artificial Life Conference Proceedings*, p. 80. https://doi.org/10.1162/isal_a_00697

100. Pigozzi, F. *et al.* (2022) Evolving modular soft robots without explicit inter-module communication using local self-attention. In *Proceedings of the Genetic and Evolutionary Computation*

101. Tang, Y. and Ha, D. (2021) The sensory neuron as a transformer: Permutation-invariant neural networks for reinforcement learning. In *Advances in Neural Information Processing Systems* (34) (Ranzato, M. and Beygelzimer, A. and Dauphin, Y. and Liang, P. and Vaughan, J.W., eds), pp. 22574–22587, Curran Associates, Inc.

102. Tang, Y. *et al.* (2020) Neuroevolution of self-interpretable agents. In *Proceedings of the 2020 Genetic and Evolutionary Computation Conference. GECCO '20*, pp. 414–424, Association for Computing Machinery, New York, NY, USA

103. Ha, D. and Tang, Y. (2022) Collective intelligence for deep learning: A survey of recent developments. *Collect. Intell.* 1. https://doi.org/10.1177/26339137221114874

104. Hafner, D. *et al.* (2025) Mastering diverse control tasks through world models. *Nature* 640, 647–653

105. Hafner, D. *et al.* (2021) Mastering atari with discrete world models. In *International Conference on Learning Representations*

106. Ha, D. and Schmidhuber, J. (2018) Recurrent world models facilitate policy evolution. In *Advances in Neural Information Processing Systems vol. 31* (Bengio, S. and Wallach, H. and Larochelle, H. and Grauman, K. and Cesa-Bianchi, N. and Garnett, R., eds), Curran Associates, Inc

107. Dhariwal, P. and Nichol, A. (2021) Diffusion models beat gans on image synthesis. In *Advances in Neural Information Processing Systems* (34) (Ranzato, M. and Beygelzimer, A. and Dauphin, Y. and Liang, P. and Vaughan, J.W., eds), pp. 8780–8794, Curran Associates, Inc.

108. Nichol, A.Q. and Dhariwal, P. (2021) Improved denoising diffusion probabilistic models. In *Proceedings of the 38th International Conference on Machine Learning (ICML)* (139), pp. 8162–8171, PMLR

109. Ho, J. *et al.* (2020) Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems* (33) (Larochelle, H. and Ranzato, M. and Hadsell, R. and Balcan, M. and Lin, H., eds), pp. 6840–6851, Curran Associates, Inc.

110. Song, J. *et al.* (2021) Denoising diffusion implicit models. In *International Conference on Learning Representations*

111. Sohl-Dickstein, J. *et al.* (2015) Deep unsupervised learning using nonequilibrium thermodynamics. In *Proceedings of the 32nd International Conference on Machine Learning vol. 37 of Proceedings of Machine Learning Research* (Bach, F. and Blei, D., eds), pp. 2256–2265, PMLR, Lille, France

112. Song, Y. *et al.* (2021) Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*

113. Rombach, R. *et al.* (2022) High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10684–10695

114. Ho, J. and Salimans, T. (2021) Classifier-free diffusion guidance. In *NeurIPS 2021 Workshop on Deep Generative Models and Downstream Applications*

115. Brooks, T. *et al.* (2024) Video generation models as world simulators. https://openai.com/research/video-generation-models-as-world-simulators

116. Jumper, J. *et al.* (2021) Highly accurate protein structure prediction with alphafold. *Nature* 596, 583–589

117. Alonso, E. *et al.* (2024) Diffusion for world modeling: Visual details matter in atari. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*

118. Zhang, Y. *et al.* (2025) Diffusion models are evolutionary algorithms. In *The Thirteenth International Conference on Learning Representations*

119. Hartl, B. *et al.* (2024) Heuristically Adaptive Diffusion-Model Evolutionary Strategy. *arXiv*, http://arxiv.org/abs/2411.13420 doi: 10.48550/arXiv.2411.13420 arXiv:2411.13420 [cs]

120. Lugmayr, A. *et al.* (2022) Repaint: Inpainting using denoising diffusion probabilistic models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11461–11471

121. Zhang, Y. *et al.* (2023) Inversion-based style transfer with diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10146–10156

122. Sclocchi, A. *et al.* (2025) A phase transition in diffusion models reveals the hierarchical nature of data. *Proc. Natl. Acad. Sci.* 122, e2408799121. https://www.pnas.org/doi/abs/10.1073/pnas.2408799121

123. Ambrogioni, L. (2025) The statistical thermodynamics of generative diffusion models: Phase transitions, symmetry breaking, and critical instability. *Entropy* 27, 291

124. Ambrogioni, L. (2024) In search of dispersed memories: Generative diffusion models are associative memory networks. *Entropy* 26, https://www.mdpi.com/1099-4300/26/5/381

125. Raya, G. and Ambrogioni, L. (2023) Spontaneous symmetry breaking in generative diffusion models. In *Advances in Neural Information Processing Systems* (36) (Oh, A. and Naumann, T. and Globerson, A. and Saenko, K. and Hardt, M. and Levine, S., eds), pp. 66377–66389, Curran Associates, Inc.

126. Hopfield, J.J. (1982) Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci.* 79, 2554–2558. https://www.pnas.org/doi/abs/10.1073/pnas.79.8.2554 doi: 10.1073/pnas.79.8.2554. arXiv: https://www.pnas.org/doi/pdf/10.1073/pnas.79.8.2554.

127. Levin, M. *et al.* (2019) Planarian regeneration as a model of anatomical homeostasis: Recent progress in biophysical and computational approaches. *Semin. Cell Dev. Biol.* 87, 125–144

128. Emmons-Bell, M. *et al.* (2015) Gap junctional blockade stochastically induces different species-specific head anatomies in genetically wild-type girardia dorotocephala flatworms. *Int. J. Mol. Sci.* 16, 27865–27896

129. Durant, F. *et al.* (2017) Long-term, stochastic editing of regenerative anatomy via targeting endogenous bioelectric gradients. *Biophys. J.* 112, 2231–2243

130. Oviedo, N.J. *et al.* (2010) Long-range neural and gap junction protein-mediated cues control polarity during planarian regeneration. *Dev. Biol.* 339, 188–199

131. Emmons-Bell, M. *et al.* (2019) Regenerative adaptation to electrochemical perturbation in planaria: A molecular analysis of physiological plasticity. *iScience* 22, 147–165

132. Csete, M. and Doyle, J. (2004) Bow ties, metabolism and disease. *Trends Biotechnol.* 22, 446–450

133. Doyle, J.C. and Csete, M. (2011) Architecture, constraints, and behavior. *Proc. Natl. Acad. Sci.* 108, 15624–15630

134. Yang, G. *et al.* (2024) Degradation and translation of maternal mrna for embryogenesis. *Trends Genet.* 40, 238–249

135. Fields, C. and Levin, M. (2018) Multiscale memory and bioelectric error correction in the cytoplasm–cytoskeleton-membrane system. *WIREs Syst. Biol. Med.* 10, e1410. https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wsbm.1410 doi: https://doi.org/10.1002/wsbm.1410. arXiv: https://wires.onlinelibrary.wiley.com/doi/pdf/10.1002/wsbm.1410

136. Bettegowda, A. (2007) Mechanisms of maternal mrna regulation: implications for mammalian early embryonic development. *Front. Biosci.* 12, 3713

137. Cummins, J. (2002) The role of maternal mitochondria during oogenesis, fertilization and embryogenesis. *Reprod. BioMed. Online* 4, 176–182

138. Bahdanau, D. *et al.* (2015) Neural machine translation by jointly learning to align and translate. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015* (Bengio, Y. and LeCun, Y., eds), Conference Track Proceedings

139. Felt, E.P. (1940) *Plant galls and gall makers*, Comstock Publishing Company, Inc, Ithaca, New York.

140. Fankhauser, G. (1945) Maintenance of normal structure in heteroploid salamander larvae, through compensation of changes in cell size by adjustment of cell number and cell shape. *J. Exp. Zool.* 100, 445–455

141. Cooke, J. (1981) Scale of body pattern adjusts to available cell number in amphibian embryos. *Nature* 290, 775–778

142. Pai, V.P. *et al.* (2012) Transmembrane voltage potential controls embryonic eye patterning in xenopus laevis. *Development* 139, 313–323

143. Durant, F. *et al.* (2019) The role of early bioelectric signals in the regeneration of planarian anterior/posterior polarity. *Biophys. J.* 116, 948–961

144. Pezzulo, G. *et al.* (2021) Bistability of somatic pattern memories: stochastic outcomes in bioelectric circuits underlying regeneration. *Philos. Trans. R. Soc. B Biol. Sci.* 376, 20190765

145. Gumuskaya, G. *et al.* (2023) Motile living biobots self-construct from adult human somatic progenitor seed cells. *Adv. Sci.* 11. https://doi.org/10.1002/advs.202303575

146. Lobo, D. *et al.* (2014) A linear-encoding model explains the variability of the target morphology in regeneration. *J. R. Soc. Interface* 11, 20130918

147. Kriegman, S. *et al.* (2021) Kinematic self-replication in reconfigurable organisms. *Proc. Natl. Acad. Sci.* 118. https://doi.org/10.1073/pnas.2112672118