



Many Paths to Anticipatory Behavior: Anticipatory Model Acquisition Across Phylogenetic and Ontogenetic Timescales

Matthew Sims¹

Received: 29 May 2022 / Accepted: 12 December 2022
© The Author(s) 2022

Abstract

Under the assumption that anticipatory models are required for anticipatory behavior, an important question arises about the different manners in which organisms acquire anticipatory models. This article aims to articulate four different non-exhaustive ways that anticipatory models might possibly be acquired over both phylogenetic and ontogenetic timescales and explore the relationships among them. To articulate these different model-acquisition mechanisms, four schematics will be introduced, each of which represents a particular acquisition structure that can be used for the purposes of comparison, analysis, and hypothesis formulation. By bringing to the fore the differences and similarities between each of the four ways that anticipatory models are acquired, a more complete picture of both anticipatory behavior and its pervasive role in biological self-maintenance can be offered. In doing so, this article helps not only to shed light on how anticipatory behavior might arise in the wide range of organisms that it has been observed in but also to throw into relief the subtle and often still overlooked causal interplay between ontogenetic and phylogenetic plasticity.

Keywords Associative learning · The Baldwin effect · Epigenetic inheritance systems · Evo-devo · Evolution · Evolution of cognition · Future-coordinated response · Model instantiation · Phenotypic plasticity

Introduction

Anticipatory behavior (henceforth AB) refers to the ability of an organism to coordinate its behavior with one or more future environmental or organismal states, proactively influencing how or whether such states occur. Such behavior has been of great interest to cognitive science and psychology because it has often been seen as indicative of cognition (Bartlett 1932; Craik 1943; Piaget 1970; Neisser 1976; Drescher 1991; Arbib 1992; Castelfranchi 2005; Lyon 2006; Pezzulo 2008; Bickhard 2009; Deacon 2012; Clark 2016; Corcoran et al. 2020; Sims 2021). Interestingly, within the last 30 years various biological studies have suggested that AB is not limited to neuronal organisms but may also be exhibited by organisms such as bacteria (Schild et al. 2007; Tagkopoulos et al. 2008), fungi (Mitchell et al. 2009; Rodaki et al. 2009), acellular slime mold (Saigusa et al. 2008), amoebae (De la Fuente et al. 2019), and plants (Novoplansky et al. 1990; Latzel and Münzbergová 2018). For example, it has been shown

that after being repeatedly exposed to episodes of three regularly spaced bursts of dry air, each separated by a 60-minute period without stimulus, acellular slime mold (*Physarum polycephalum*) spontaneously decreases its locomotion speed prior to the onset of when the next air burst episode would have occurred were the same temporal pattern continued (Saigusa et al. 2008). Whether this example or any of the other instances of AB in non-neuronal organisms should truly be considered an expression of cognition is a matter of contested debate (see Corcoran et al. 2020). Setting the topic of cognition aside, one important question that arises for AB across the biological board is how its various instances arise in the first place. That is, what imbues an organism, whether an *E. coli* or a human, with the capacity to coordinate its behavior with yet-to-be encountered environmental (or organismal) states?

One answer that has been suggested by numerous scholars is that AB requires that organisms instantiate internal anticipatory models, the structure of which simultaneously captures the structure of the organism's environment and the manner in which the organism's actions affect both it and its environment (Rosen 1985/2011; Tagkopoulos et al. 2008; Friston et al. 2010, 2019; Louie 2010, 2012; Ginsburg and Jablonka 2019; Levin 2019; Corcoran et al. 2020; Sims 2021; Sims and Pezzulo 2021). In acquiring an anticipatory model, organisms may exploit such

✉ Matthew Sims
matthew.sims-m4e@RUB.de

¹ Institute for Philosophy II, Ruhr-University Bochum, Bochum, Germany

models to guide anticipatory adaptive control, preparing for resource fluctuations or mounting defensive responses prior to encountering a threat. If it is indeed the case that anticipatory models are required for AB, an important question arises about the different manners in which organisms acquire anticipatory models.

This article aims to articulate four different non-exhaustive ways that anticipatory models might possibly be acquired over both phylogenetic and ontogenetic timescales and explore the relationships between them. They are: (1) mutation-based acquisition, (2) epigenetic inheritance-based acquisition, (3) associative learning-based acquisition, and (4) Baldwin effect-based acquisition. To articulate these different model acquisition mechanisms, four schematics shall be introduced, each of which represents a particular acquisition structure that can be used for the purposes of comparison, analysis, and hypothesis formulation. It is my hope that by bringing to the fore the differences and similarities between each of the four ways that anticipatory models are acquired, a more complete picture of both AB and its pervasive role in biological self-maintenance can be offered. In doing so, this article helps shed light on how AB might arise in the wide range of organisms that it has been observed in, something that is important under the assumption that not every form of model acquisition is found across all taxa. Associative learning-based acquisition, for example, might go a long way in explaining how many cases of AB arise in many animals (Ginsburg and Jablonka 2019) and plants (Gagliano et al. 2016; but see Markel 2020) and amoebae (De la Fuente et al. 2019; Carrasco-Pujante et al. 2021); however, it fails to account for AB in bacteria, yeast, or slime mold, organisms for which there is currently a lack of substantiated evidence supporting the presence of associative learning. Something else in these organisms must account for their AB. By exploring some of the relationships among these possible paths to AB and hence throwing into relief the subtle and often still-overlooked causal interplay between ontogenetic and phylogenetic plasticity, this article moves the discussion forward about the relationship between learning and evolution, a foundational topic in both biology and cognitive science.

The organization of this article proceeds as follows: the second section provides a brief overview of the notion of anticipatory models and their relation to AB as proposed by Robert Rosen (1985/2011). The third section articulates four possible ways that anticipatory models might be acquired, proposing four respective core mechanisms, and providing a simplified schematic for each. To this, examples supporting the presence of each mechanism, with the exception of associative learning-based acquisition, shall be provided. Taking into consideration the characterization of the four manners of acquiring anticipatory models, the fourth section considers some of the environmental and organismal features that constrain each acquisition type, thereby offering a few rough heuristics to guide both the identification of the different model-acquisition types in nature and experimental evolutionary design.

Before moving on, a few preliminary remarks are in order: the notion of behavior, as it will be used throughout this article, refers broadly to both motor responses (i.e., movement or non-movement produced by striated contractile tissues) and to biochemical responses (i.e., the production or inhibition of biochemical gene products) at the level of an individual organism. This notion of behavior diverges significantly from that deployed in comparative psychology and is more akin to how behavior is understood in the process theory, active inference (Friston et al. 2017), and within biogenic approaches to cognition (Lyon 2006) such as autopoietic theory (Maturana and Varela 1980).

Moreover, in this article I shall follow Rosen (1985/2011) in assuming that all AB requires that a system instantiate an anticipatory model (also see Poli 2010). Just how such models may be construed is the topic of the section to which we now turn.

A General Overview of Anticipatory Models

The idea that AB results from the exploitation of anticipatory models is one that has been extensively developed and defended in the context of systems biology by theoretical biologist Robert Rosen (1985/2011). According to Rosen, living systems are *anticipatory systems*. Their capacity to adaptively coordinate their behavior with the future is something that Rosen argued sets living systems apart from nonliving systems. An anticipatory system is one that “contains a predictive model of itself and/or of its environment, which allows it to change state at an instant in accord with the model’s predictions pertaining to a latter instant” (Rosen 1985/2011, p. 313). If living systems are anticipatory systems, then, as this characterization makes clear, the concept of an anticipatory model is central to explaining any instance of AB across the range of all living systems. More recently, anticipatory models have found a statistical analogue in the notion of recognition models within the theory of active inference (Friston 2010; Pezzulo et al. 2015; Corcoran et al. 2020; Tschantz et al. 2020) where they, together with generative models, function to proactively control perception and action (Sims and Pezzulo 2021). Avoiding unnecessary technical details, in this section I will provide a general overview of anticipatory models. The notion of anticipatory models that shall be presented here is not meant to be faithful to Rosen’s account or to active inference but may be seen as largely compatible with such accounts of anticipatory models. The purpose of this overview is to get a clear general picture of what such internal models are, a picture that can be referred to later when considering the various ways that I will suggest anticipatory models might be acquired.

In order to grasp what anticipatory models are, it is first necessary to understand what they are models of and hence the conditions that make anticipatory models possible. Living systems inhabit environments in which such *regularly sequenced structural relations* are ubiquitous (Betchel 2011; Feddilino and Tavazoie 2012; Bernhardt et al. 2020). To be sure, correlational structures may take at least two forms: one form describing the structure that exists between an environmental state/event type at one time and that same state/event's tendency to follow a particular trajectory over time. The example above of periodic dry air bursts that *Physarum polycephalum* proactively responds to is an illustration of this form of correlational structure. Another form of correlational structure describes the relation between one environmental event type that follows regularly from another environmental event type. It is this latter kind of correlational structure that the current article will be focused upon. Whether it is the tendency that an increase in temperature will be followed by a decrease in ambient oxygen upon entering the mammalian gastrointestinal tract (see below) or the tendency for fluid remains of the freshwater crustation (daphnia) to be followed by the continued presence of their predators in the next generation (Laforsch et al. 2006), such regular temporal correlations abound. Importantly, these structures between events can occur across many timescales, ranging from one event type following another in a matter of seconds to being separated by one or more generations. The correlated events that this article is concerned with are those that occur within the lifetime of an individual organism at least once. As we shall see, it is often the case that the temporal nature of a given correlational structure determines how that structure is captured in the dynamics of an anticipatory model. It is the existence of these regular structures *and* the fact that they can be both harnessed and exploited for adaptive behavior that the notion of anticipatory models is meant to capture.

An anticipatory model may be characterized as a living system's way of being organized, such that the organism's dynamic patterns of organization capture the regular correlational structure of the environment and/or organismal sensory state trajectories (Bernhardt et al. 2020; Sims 2021) (see also Landmann et al. 2021 for "model predictive control").¹ Capturing the environmental and/or organismal structure is a process of abstraction (Rosen 1985/2011) and

refers to the idea that a subset of states (properties or characteristics) of the modelled system are mapped *onto* some set of distinct states (properties or characteristics) of the modelling organism. It is because this mapping relation is *structure preserving* that state changes in the model can be reliably correlated with state changes in the environment and/or sensory states.² Structured model dynamics, as a result, may be used to drive behavior that is adapted to environmental structure prior to encountering part of the structure.

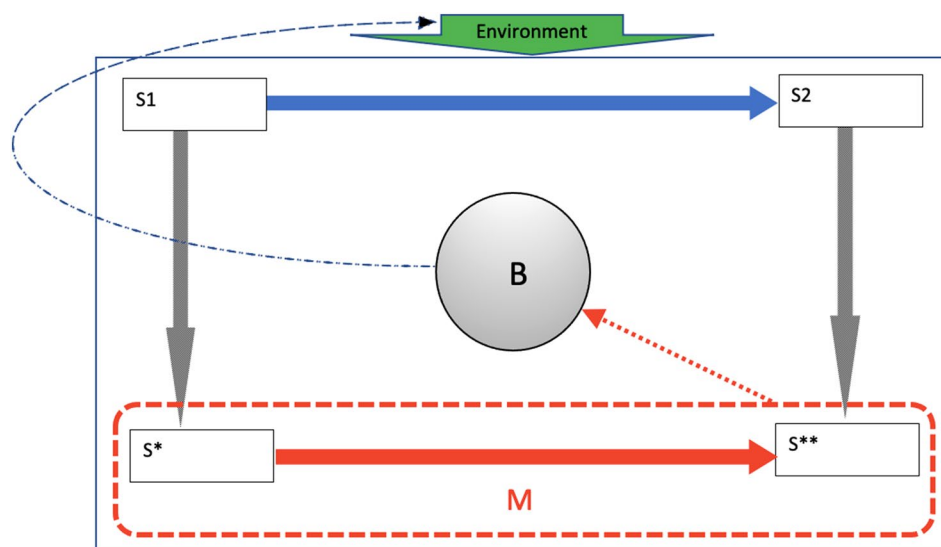
Importantly, anticipatory models do not capture just any property or characteristic of the target system. They capture regular correlational structure. Such capturing may be cast as a form of data compression in which many instances of observation pairs are encoded into a compact rule (i.e., a generalization). It should be noted here that model deployment and rule-following are subpersonal-level process that need not (but can at some level of increased complexity) rise to personal-level awareness (Poli 2010; Clark 2016; Corcoran et al. 2020; Sims 2021). Crucially, it is because the model's structure-capturing dynamics are faster than the dynamics of the modelled system that the outcome state in the model is anticipatory of the outcome state of the modelled system/process (Rosen 1985/2011; Pezzulo 2008; Poli 2009; Louie 2010).

For example, say there is an organism, *O*, with a model, *M*. Let us also say that *M* captures a structural relation between two sensory states, call them *S1* and *S2*, such that *S1* maps onto *S** and *S2* maps onto *S*** in *M*. The relation between *S1* and *S2* is such that the latter regularly follows the former and in *M* this maps onto [if *S** at t_n then *S*** at t_{n+1}]. In other words, the fact that *S** occurs prior to *S*** in the model encodes information about the temporal order of the environmental event types that cause sensory states *S1* and *S2*. If state *S1* at t_1 is encountered, then because of the captured correlational relation and the fact that trajectory of *M*'s dynamics (going from *S** to *S***) runs faster than the trajectory of the environmental states causing both *S1* and *S2* and hence faster than *S1* to *S2*, *M* can elicit predictive dynamics, allowing *O* to behaviorally cope with *S2* upon encountering *S**. That is, as a result of the mapping of states and structural relations in *M*, *M*'s dynamics can be exploited by *O* to anticipate an environmental/sensory state at a future

¹ Interestingly Riegler (2001) writes: "Can we find an answer to the question of whether anticipation needs an internal model in the mind or whether it is a fundamentally embedded in the organization of the subject?" (p. 1). If the characterization I've provided of anticipatory models here is accurate, then posing such a question introduces a false dichotomy; internal models *are* fundamentally embedded in the organization of the subject.

² In Rosen's theory, the notion of structural preservation is spelled out mathematically in terms of *conjugacy* between sets of states in a natural system and sets of states in a model. For the details see Rosen (1985/2011). In active inference, structural preservation is given a statistical gloss as the tuning of the priors that feature in an internal model (see Sims and Pezzulo 2021).

Fig. 1 Schematic of an anticipatory model: S1 and S2=regularly successive sensory states; S* and S**= successive organismal states; M= the model; B=behavior



timestep and proactively behave, B, in a way that prepares it for S2 prior to S2's coming about (see Fig. 1).

The benefit of a cue to an organism rests in its ability to reduce environmental uncertainty about future states in virtue of the cue and the future state exhibiting high mutual information (Donaldson-Matasci et al. 2010; Bernhart et al. 2020).³ Given that subsequent environmental states that follow the presence cue (S1) are often non-neutral with respect to how such states affect the organism, exploiting anticipatory models is key for the control of adaptive behavior. For example, hearing rapid and evolving rhythmic patterns against the surface of a stairway cupola is not only a cue that it is raining outside but also that your plans to walk to the shop now will most likely result in your getting drenched. By capturing correlational structure between the occurrence of certain environmental states and the environmental and sensory outcomes that follow from such states, a model may be leveraged to generate behavior that influences whether or how those outcomes unfold (walk to the shop later if you prefer not to get wet!). This is particularly important when the negative effects that might follow the occurrence of a particular environmental (cue) state threaten to destabilize long-term homeostasis. Given that homeostatic equilibrium is dependent upon an organism's remaining in a select range of phenotypic states (its homeostatic range) despite environmental perturbation (Cannon 1932), AB is driven and nuanced by anticipatory models that fundamentally answer to an organism's long-term homeostasis.⁴ It

is via future-oriented internal dynamics in the model that an organism's behavior is tuned so as to continue to bring about homeostasis-compatible sensory states (Sims 2021). Thus, AB—when all goes well—may make the difference between something as menial as staying dry and getting wet but it also may be that which makes the difference between defending or foregoing biological integrity.

Hence, we arrive at Rosen's (1985/2011) claim that living systems are anticipatory systems; it is difficult to imagine purely responsive systems would live very long given that such a system's behavior would be error driven. Such systems would be condemned to a (short) life of corrective responses and since it takes time to initiate a corrective response, they would run the risk of irrecoverable dyshomeostasis prior to recovery (cf. Poli 2009, 2010; Bernhart et al. 2020; Sims 2021). As such, we should expect anticipatory models that allow for proactive rather than reactive behavior to be a ubiquitous feature of the living world (Louie 2010). It is in virtue of its contribution to the robustness at the level of the individual that "anticipation also becomes one of the primary drivers in the process of evolution and adaptation as well of a product of it" (Rosen 2009, pp. 8–9).

With this general notion of anticipatory model to hand, let us now move onto exploring four possible ways that anticipatory models might be acquired across a diverse range of organisms.

Four Paths to Anticipatory Behavior

In this section I will begin by articulating some of the simpler mechanisms that underwrite anticipatory model

³ Mutual information is a probabilistic measure that quantifies the mutual dependence between two random variables when sampled simultaneously.

⁴ The homeostatic range that defines an organism's phenotype is interpreted in active inference in terms of its generative model, a stipulated normative model that is entailed by the realized (anticipatory) recognition model.

acquisition; given the simplicity of these proposed mechanisms it is reasonable to think that they are good contenders for anticipatory model acquisition in basal organisms. I will then proceed to work my way up, looking at more complex mechanisms that may or may not be restricted to explaining anticipatory model acquisition in more complex organisms. In presenting these mechanisms from simple to complex, the evolutionary order in which such mechanisms may have historically arisen is hinted at. The various paths to AB that I will trace out here span a range from phylogenetic acquisition to ontogenetic acquisition and each successive mechanism presented builds upon the details of the mechanism(s) that preceded it. The nature of this section involves some *empirically supported* speculation. However, as long as the content that is core to any speculation is subject to adjudication and is able to inform testable hypotheses, then speculation itself remains an invaluable part of scientific discourse. As such, nothing that will be presented in this section is meant to decide anything a priori. It should merely serve as a call to and a tool for further investigation theoretically, computationally, and at the bench.

Mutation-Based Acquisition

A first possible manner that organisms might acquire an anticipatory model that I would like to explore is based on genetic variation via random mutation. Mutation-based model acquisition refers roughly to the idea that some subpopulation of organisms comes to capture the correlational structure of the environment after repeatedly encountering that structure over multiple generations via the role that heritable random mutation (e.g., a permanent modification in the DNA sequence of a gene) plays in an embedding regulatory network. The presence of the mutation—as interpreted—results in an otherwise merely reactive regulatory network exhibiting anticipatory dynamics (i.e., instantiating an anticipatory model) and the form that these dynamics take may be transmitted across generations.

In an extremely simplified representation of mutation-based acquisition we might imagine something like the following occurring: members of a microbial subpopulation encounter some environmental cue that causes a sensory state (call it S_1).⁵ As a result of that encounter, the gene

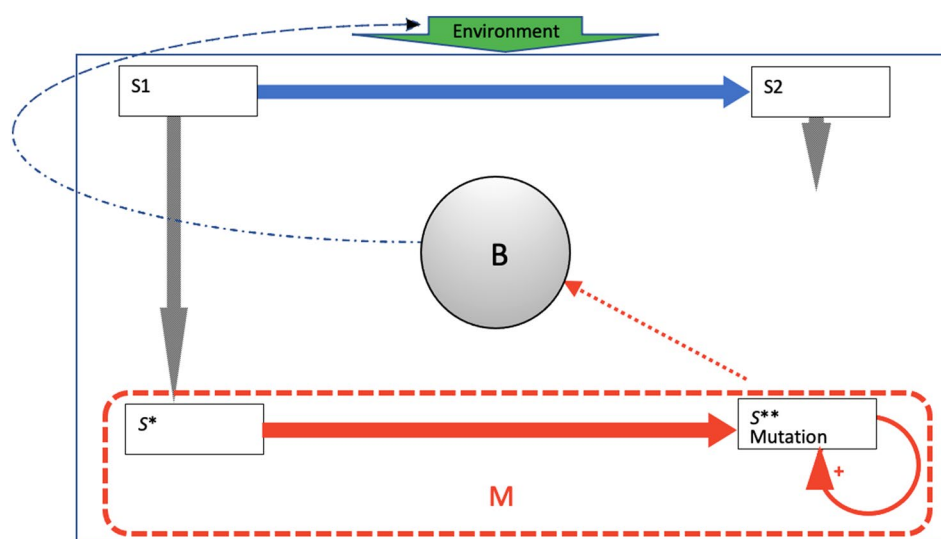
product that is expressed (or repressed) (call it S^*) in the organism sets into motion the expression of a secondary, *auto-catalyzing* gene product (call it S^{**}) in virtue of the mutation of gene S^{**} . Autocatalysis is a kind of positive feedback mechanism in which the presence of the auto-catalyzing product both causes and speeds up its own production and as such is produced rapidly in a short amount of time (cf. Kauffman 2000). Given the specificity of the environmental inducer of S^* , and that S^{**} can be induced by either gene product S^* or by the environment via S_2 , the configuration between genes S^* and S^{**} encodes information about the regular temporal order of the environmental event types that cause S_1 and S_2 . As such, the induction of S^* raises the probability that S_2 will occur afterwards. If the product S^{**} is both produced more rapidly than the environmental state trajectory causing both S_1 and the subsequent sensory state (S_2) and the expression of S^{**} allows the organism to compensate for the yet-to-be encountered environmental cause of S_2 to the degree that without S^{**} the probability of a cell's surviving would decrease, then mutation-based acquisition can occur. This compensation takes the form of some anticipatory (meta-metabolic, physiological, or motoric) behavior (B). Importantly, in those organisms lacking the gene S^{**} mutation, the gene product S^* does not induce S^{**} ; rather, product S^{**} is induced directly by sensory state S_2 (i.e., by encountering the environmental change that causes S_2) (see Fig. 2).

Given the presence of the mutation and its effect on the regulatory pathway and the regular correlational structure in the environment, as long as the conditions brought about in the latter act as a stable selection pressure for the population, and there happens to be no other mechanism that the population can deploy more efficiently to cope with that selection pressure, the genetic variation will confer a fitness advantage to the subpopulation. That is, via natural selection, increased differential fitness will result in an eventual increase in the frequency of the random mutation within the population until most of the population acquires the anticipatory model, or the niche is altered in such a way that the variation is no longer adaptive.

Since natural selection is the driving mechanism behind this way of acquiring anticipatory models, and because natural selection when driven alone by random mutation is a slow and gradual process, mutation-based acquisition is itself slow and gradual. That being said, for organisms such as bacteria, because of their capacities to reproduce sexually via conjugation in addition to both picking up genetic material directly from their environments and merging via fusion, the time it takes for such mutation to propagate through a bacterial population is comparatively much less than it is for organisms that are limited to sexual reproduction alone for the exchange of genetic material. This is not a trivial point: once an individual bacterium inherits this form of

⁵ The simplified schematics I offer throughout this article, like all other models, abstract away from the many details that we might expect to find in the target phenomena. For example, as it may occur in actual organisms, mutation-based acquisition will certainly involve complex gene regulatory network dynamics in which multiple genes and proteins interact by way of transcription factors and RNA. Neither these complex regulatory interactions nor the constraints put in place by epistasis are featured in the mutation-based acquisition schematic.

Fig. 2 Schematic of mutation-based anticipatory acquisition: S1 and S2=regularly successive sensory states; S^* =gene; S^{**} =gene mutation with auto-catalyzing gene product; M=the model; B=behavior. S1 induces the production of S^* which leads to the expression of secondary auto-catalyzing gene product S^{**} . Given the high rate at which S^{**} is produced and its influence on B, the regulatory network dynamics that embody M result in proactive B prior to S2



mutation-based anticipatory model, the fast-paced rate of asexual reproduction (binary fission) is likely to ensure that all daughter cells, having an identical genome, carry the mutation and hence the model.

Under what conditions might mutation-based model acquisition arise? Generally, such model acquisition would be likely to occur in instances in which populations face new selective pressures because of their exposure to new environmental conditions that are potentially dyshomeostasis inducing. Such exposure may be non-exhaustively driven by (a) a population's colonizing a new environment that exhibits an already present exploitable regular structure; (b) a population modifying their own environment over time;⁶ or (c) the colonization of an already occupied environment by new populations (i.e., predators and/or competitors), and those colonizers introducing new exploitable structures for those organisms that were already there. Given the complex nature of ecologies (Gilbert 2000; Watson and Szathmary 2016), it is most likely that all three of these conditions set the stage for mutation-based acquisition to arise.

That mutation-based acquisition is a plausible explanation for at least some of the AB exhibited by basal organisms finds support in studies (and the interpretations thereof) on AB in bacteria (Tagkopoulos et al. 2008; Mitchell et al. 2009; Freddolino and Tavazoie 2012) and yeast cells (Mitchell et al. 2009; Freddolino and Tavazoie 2012; Dhar et al. 2013; Mitchell and Lim 2016). For instance, studies by Tagkopoulos et al. (2008) provided evidence for anticipatory behavior in *E. coli* which occurs in the metabolically

important transition from aerobic to anaerobic environments, the kind of transition that *E. coli* encounter when migrating from a terrestrial environment into the gastrointestinal tract of a mammal. Using a bioreactor to simulate the kind of correlational structure between a temperature upshift (25 °C to 37 °C) and an ambient oxygen downshift that would be encountered as a bacterium enters the GI tract, Tagkopoulos et al. (2008) found that *E. coli* use the temperature upshift as cue to both repress genes encoding for products that allow for aerobic respiration (cytochrome bo3 oxidase complex and components of the TCA cycle) and induce genes that allow them to engage in anaerobic respiration (components of the cytochrome bd oxidase complex) prior to encountering an oxygen-poor environment. Redescribing these findings using the simplified schematic of mutation-based acquisition: a sensed temperature upshift (S1) results in the repression of aerobic respiration-enabling gene products (S^*), which induces the expression of anaerobic respiration-enabling gene products (S^{**}), the activity of S^* and S^{**} together engaging an anaerobic metabolic system prior to the bacterium's sensing oxygen-poor environmental conditions (S2).

The coregulation of genes mediating such anticipatory bacterial behavior, according to Tagkopoulos et al., is an adaptive anticipatory response to the correlational structure between temperature change and oxygen change that has evolved over phylogenetic timescales. It is clear that these researchers see something like mutation-based model acquisition as being responsible for the presence of *E. coli*'s anticipatory response as evidenced when they write:

More generally, the correlation-structure of the environment can be internalized as a probabilistic model in the high-dimensional space of an organism's complete

⁶ For example, the presence of oxygenated environments on Earth was preceded by an anaerobic environment populated with bacteria whose metabolic waste created the aerobic environment that they had to metabolically adapt to.

sensory perception. As such, the very organization of microbial regulatory networks may, in large part, represent the physical instantiation of this probabilistic model. (Tagkopoulos et al. 2008, p. 6)⁷

The idea that AB is an evolved response and that the capacity for AB in microorganisms is underwritten by an internalized anticipatory model of the environment's correlational structure, suggests that the acquisition of an anticipatory model in the case of *E. coli*'s anticipatory metabolic regulation may indeed be explained in a manner that is consistent with mutation-based model acquisition (see also Mitchell and Lim 2016; Freddolino and Tavazoie 2012).⁸ The plausibility of this interpretation for these striking experimental results is dependent upon the assumption that heritable variation takes the form of gene variation (i.e., classical mutation). We, however, will see in the next subsection that this needn't always be the case.

Epigenetic Inheritance-Based Acquisition

Before articulating the details of the next form of model acquisition that I would like to propose, a few remarks are in order regarding the notion of epigenetic inheritance itself. This form of inheritance refers to “the transmission to subsequent generations of cells or organisms of phenotypic variations that do not stem from variations in the DNA base sequence” (Jablonka 2017, p. 3). As such, it represents a form of *soft inheritance*: a case in which the environment can directly influence variation in heritable material, having persistent, adaptive phenotypic effects. Such epigenetic variation may be induced by the external environment, internal regulatory factors, and random fluctuations in the cellular milieu (Jablonka and Lamb 2005).⁹ Although the molecular mechanisms underwriting epigenetic inheritance are still not known in many systems, some of the mechanisms that are currently recognized include *structural templating* (e.g., conformation of prion proteins) (Harvey et al. 2020),

self-sustaining gene regulatory loops (Roberts and Wickner 2003; Dubnau and Losick 2006; Zordan et al. 2006), *small RNA interference* (Nowacki et al. 2008; Rechavi et al. 2011), and *chromatin marking* (e.g., patterns of DNA methylation and histone modification) (Chandler 2007; Castel and Martienssen 2013).¹⁰ Importantly, these mechanisms are overlapping and interdependent (Cedar and Bergman 2009; Bateson and Gluckman 2011), the interaction between small noncoding RNAs and other epigenetic mechanisms likely playing a major role in providing specificity to other epigenetic mechanisms (Kozioł and Rinn 2010).

Transmission of phenotypic variations can occur through asexual reproduction (i.e., binary fission in prokaryotes) (Roberts and Wickner 2003; Dubnau and Losick 2006; Hu et al. 2018), in addition to meiotic (germ cell) division (Champagne and Meaney 2006; Rechavi et al. 2011; Rodgers et al. 2013; Yu et al. 2013).¹¹ Through the transmission of functionally adaptive phenotypic states, offspring may be *preadapted* to some of the environmental conditions that their parent(s) encountered, hence avoiding the cost of deleterious effects that accompany being in nonadaptive phenotypic states and/or the cost of having to expend metabolic energy and time on switching phenotypes.¹² As such, epigenetic inheritance can contribute to an individual's selective advantage, facilitating evolutionary change as a source of heritable variation. Moreover, it can also increase the rate of evolution because epigenetic variation can occur simultaneously

⁷ To describe a model as being probabilistic is to say that the correlational rules that govern the state transitions in a model (and hence behavior) are governed by something akin to Bayes's rule. We have already seen above that active inference (Friston et al. 2017; Kiefer and Hohwy 2018; Tschantz et al. 2020) is committed to a similar statistical understanding of anticipatory model dynamics.

⁸ Circadian rhythms may be plausibly understood as a kind of instantiated anticipatory model (Tagkopoulos et al. 2008), one that may have been acquired early on in evolutionary history via mutation-based acquisition in ancient photosynthetic bacteria such as Cyanobacteria (see Edgar et al. 2012 for a detailed theory on the changing environmental conditions that drove the evolution of circadian rhythms).

⁹ Strictly speaking, epigenetically inherited phenotypes can also be the product of random differential DNA methylation regions, or what are known as *epimutations* (see Holliday 1987).

¹⁰ For a detailed overview of these mechanisms see Jablonka and Raz (2008) and Heard and Martienssen (2014).

¹¹ Although there is evidence suggesting that histone modification may play a role in mediating transgenerational epigenetic inheritance in invertebrates and plants (Lim and Brunet 2013; Heard and Martienssen 2014; Fallet et al. 2020), the notion of transgenerational epigenetic inheritance remains controversial (Radman-Livaja et al. 2010; Ptashne 2012; Cantone et al. 2013). One reason for such skepticism is that histone marks (which include histone methylation, acetylation, and phosphorylation) are removed during mammalian oogenesis. However, while such marks are also removed during spermatogenesis in mammals, they are replaced by protamines which may play a role in transgenerational epigenetic inheritance in mammals (Brunner et al. 2014). DNA methylation marks, on the other hand, are usually removed in both males and females in mammals but may be reconstituted in the next generation, presumably through interactions with transmitted gametic RNAs (these two epigenetic mechanisms are often linked, it seems) (Grewal and Jia 2007; Ringrose and Paro 2007). Moreover, there seem to be many genomic regions in vertebrates that may not be subject to demethylation (see Van der Heijden et al. 2006), thus warranting further investigation into DNA methylation as a mediator of transgenerational epigenetic inheritance in vertebrates (Nestler 2016).

¹² In the sense that offspring are preadapted, epigenetic inheritance may be cast as a developmental anticipatory process at the level of the lineage. Although interesting in its own right, I would like to remind the reader that here I am interested in AB exclusively at the level of the individual.

in multiple individuals when the inducing environmental conditions are encountered (Jablonka and Lamb 2020). Importantly, the kinds of inherited responses that are adaptive are those that “can be inherited for a while in the absence of the environmental trigger, but not for a very long period of time” (Lachmann and Jablonka 1996, p. 5). Progeny is thus preadapted, and yet over the course of ontogeny if environmental conditions change and such phenotypic states are no longer adaptive in the current context, they may be (developmentally) selected against whilst other heritable adaptive phenotypic variations are environmentally induced (cf. Soen et al. 2015).¹³ Epigenetic inheritance thus represents a significant form of phenotypic plasticity.

Anticipatory model acquisition that is based upon epigenetic inheritance, I would like to suggest, refers to the idea that heritable epigenetic variation via any of the aforementioned epigenetic mechanisms can capture environmental correlational structures that organisms are regularly exposed to over time and thus lead to adaptive anticipatory behavioral phenotypes that may be transmitted across generations. Unlike mutation-based acquisition, this form of model acquisition does not have to wait upon the presence of a random mutation to occur. Like mutation-based acquisition, however, the regularity of a particular correlational structure in the environment at the level of epigenetic inheritance-based acquisition reinforces a model acquisition *across generations*: in order for the model to be acquired, it is not the case that a single individual must encounter the correlational structure multiple times in its lifetime; rather, that structure needs only to be regularly encountered by the population over many generations. Moreover, given that epigenetic inheritance mechanisms such as chromatin marking, self-sustaining loops, and regulation and heritability of small RNAs are taxonomically ubiquitous (Ginsburg and Jablonka 2009), anticipatory model acquisition based on these particular mechanisms may also underwrite the capacity for AB in a host of organisms of varying complexity—from prokaryotes to eukaryotes. For reasons of limited space my current focus will however be restricted to offering a proposal as to how epigenetic inheritance-based acquisition might occur in virtue of a *general epigenetic mechanism*—a mechanism that can be filled out more accurately with

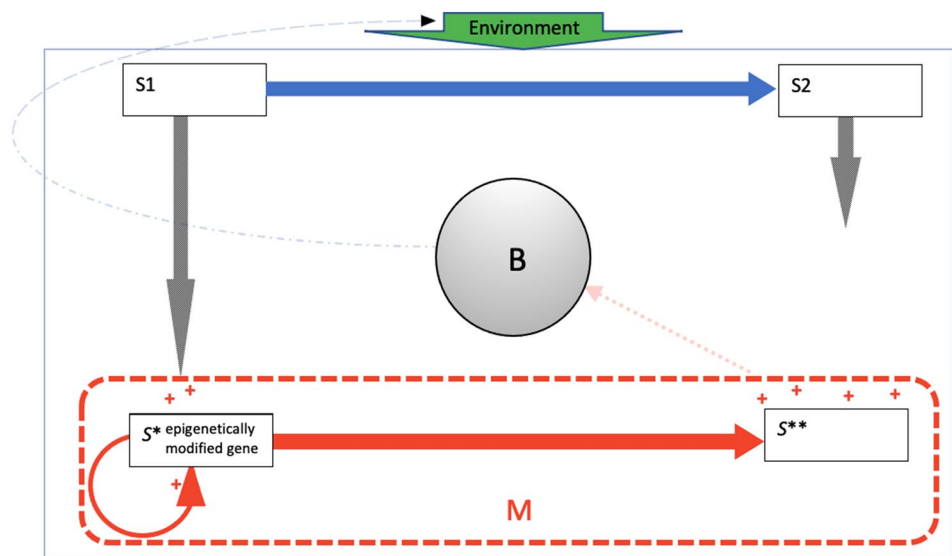
details of chromatin marking, self-sustaining loops, and/or small RNA interference.

For the basis for the schematic, let’s consider a hypothetical asexually reproducing organism. This organism encounters a change in its environmental conditions via sensory state S1 for the first time. Assume also that the environmental cause of S1 is regularly followed by (in a manner of minutes) a stress-inducing event, the cause of S2. The continued presence of S1 induces the expression of gene product S*, the expression of which is otherwise normally inhibited in conspecifics lacking any encounters with environmental causes of S1. The occasioning of S1 also results in the epigenetic modification of gene S*, engendering autocatalysis with respect to product S*. As long as the signal from S1 is occasioned, it amplifies the production of S*. When S* has reached a high enough concentration, it affects gene S** by acting itself as an amplifier of cellular signal S2 at gene S** and thereby lowering the threshold of the specific signal associated with S2 required to induce the expression of product S**. Such amplification, for instance, may take the form of the binding of multiple copies of a transcription factor to a promotor. Importantly, the concentration level of S* required for the maintenance of the autocatalytic loop is not as high as the level required for amplifying the signal S2 at gene S**. When and only when the environmental cause of S2 is encountered—even at low trace values—is the expression of S** induced, allowing the organism to adapt to homeostatically challenging environmental states prior to encountering them by making a difference as to how a resulting biochemical behavioral response (B) will come about. More precisely, the activity of the epigenetic mechanism as situated within the dynamics of the regulatory network *primes biochemical behavior*, allowing for a more rapid and stronger response to S2 when (and if) it arises. The epigenetically modified activity of gene S* in relation to the disposition of gene S** to be affected by S* (i.e., the anticipatory model) respectively encodes information about the fact that a specific environmental event type that uniquely engenders S1 has been detected by the organism in its lifetime (or in the lifetime of the organism’s progenitor(s)) and how the temporal order of the environmental event types that uniquely engender S1 and S2 are likely to come about in a natural environment: there a high probability that S1 will be followed by S2 (see Fig. 3).

When S1 is no longer occasioned, the levels of S* slowly decrease, but because of the autocatalytic activity engendered by the epigenetic mechanism, product S* remains present in the network. Although the concentration of S* does not remain high enough to have effects upon gene S** (and thus the occasioning of S2 without S1 will not elicit a primed response), it remains high enough to self-perpetuate at gene S*. Once the epigenetic molecular mechanism is in place and some form of autocatalysis is occurring, given that the

¹³ The idea that phenotypic variation in epigenetic inheritance can be environmentally induced seems to be very similar to the general notion of phenotypic accommodation (Baldwin 1896, 1902; West-Eberhard 2003). Although related, there are important differences between the two notions. In phenotypic accommodation, there is environmentally inducible phenotypic variation *without* heritability or genetic change. In epigenetic inheritance, however, although environmentally induced phenotypic variation is not a genetic change, it is heritable via epigenetic mechanisms. Moreover, it is heritable despite the absence of the environmental state that originally induced the phenotype.

Fig. 3 Schematic of epigenetic inheritance-based model anticipatory acquisition: S_1 and S_2 = regularly successive sensory states; S^* = epigenetically modified gene; S^{**} = gene with lowered induction threshold; M = the model; B = behavior. S_1 induces epigenetic modification of S^* , the auto-catalyzing product of which reduces the threshold of reaction of S^{**} to input from S_2 . M consists of the epigenetically mediated regulatory dynamics between S^* and S^{**} which brings about a more rapid B in response to S_2



markers involved in such a mechanism can be transmitted from parent to offspring, the network dynamics between the autocatalyzing S^* and gene S^{**} (i.e., the model M) can be transmitted across generations. In short, anticipatory models can be maintained across generations in virtue of the fact that epigenetic markers act as a type of transgenerational memory *and* that the configuration between S^* and S^{**} has already been genetically encoded and may thus be inherited.¹⁴ If S_1 is occasioned in the same organism or its progeny, levels of S^* will be amplified (again) to the degree that this product will prime gene S^{**} for a stronger and faster response to the highly probable occurrence of S_2 . The adaptive advantage of this form of model acquisition lies in the fact that after reproduction (e.g., division or budding) it is more physiologically/metabolically efficient for offspring to be preadapted to the environment rather than having to endure a period of time adapting to it—particularly if the environmental correlational structure encountered by the offspring is similar to that which was experienced by the parent.

Since some forms of epigenetic variations are heritable, the phenotypes they engender can be acted upon by natural selection, such that those members of the population exhibiting epigenetic marks that enable AB, given the continuation

of the environment's correlational structure, will have a selective advantage over those organisms that fail to exhibit such nongenetic variations. This proposed manner in which anticipatory models might spread throughout a population over generations exemplifies an important feature of both epigenetic and genetic inheritance; namely that in natural circumstances both forms of inheritance exert an influence upon one another. That said, since epigenetic variations are both “context-sensitive and more frequent than genetic variations, they may often initiate, bias, and facilitate evolutionary change” (Jablonka 2017, p. 6). For example, by allowing fast adaptation to correlational structure in the environment, epigenetic inheritance-based acquisition allows organisms to be situated in a niche where certain mutations can be advantageous and some of these mutations might indeed allow mutation-based acquisition to get a foothold.¹⁵

One phenomenon that lends support to the idea that epigenetic inheritance-based model acquisition is found in the natural world is transgenerational priming in plants (Sobral et al. 2021). Depending upon the nature of their encounter with herbivores, plants' defense metabolite expression can be either *directly induced* or *indirectly induced* to increase their resistance to herbivore damage (Morrell and Kessler 2014). Direct induction occurs when a plant induces the expression of defense metabolites as a result of the detection of herbivore salivatory chemicals. Indirect induction is a form of priming in which a plant detects volatile organic compounds (VOCs) emitted by neighboring herbivore-damaged plants and these VOCs act as cues, *priming* the plant for both faster and stronger defense responses (direct induction) in

¹⁴ If the epigenetic mark is somehow erased over meiosis, the model is lost despite the configuration between genes S^* and S^{**} being kept intact. Similarly, if a random mutation disturbs the configuration between genes S^* and S^{**} , then the model is lost despite the maintenance of the epigenetic mark. In this sense, anticipatory models acquired via epigenetic inheritance build upon the correlational structure that has already been captured by gene regulatory networks, but neither can such models nor their acquisition be reduced to the structure that has been captured by gene regulatory networks over phylogenetic timescales.

¹⁵ For possible examples of how the epigenome may influence mutation see Anastasiadi and Piferrer (2019) and Monroe et al. (2022).

anticipation of the likely future encounters with herbivores (Morrell and Kessler 2014). I would like to argue that when indirect induction (i.e., priming) is taken into consideration with two additional facts, a clearer picture of evidence for epigenetic inheritance-based model acquisition starts to reveal itself.

The first fact is that VOC priming is thought to be underwritten by epigenetic molecular mechanisms (Kim and Felton 2013; Morrell and Kessler 2014). The second is that herbivore-related induction in mother plants can lead to increased defense responses in offspring, and evidence suggests that this transgenerational priming capacity is the result of offspring inheriting the methylated states of herbivore-sensitive loci from their mothers (Sobral et al. 2021). These authors, commenting on the results of their study on the wild radish (*Raphanus sativus*), write:

This work shows that the patterns of defense deployment through plant ontogeny are partly shaped by herbivore-induced plasticity not only in the current generation but across generations too. Such influence of herbivory in the former generation can be directly expressed (direct induction) or retained as priming, a hidden potential, and thus expressed only in response to the appropriate herbivore cues early or late in the plant's life. (Sobral et al. 2021, p. 2)

Together these facts suggest that the reconfigured nucleosome and its subsequent activity within a regulatory network instantiates an anticipatory model acquired via epigenetic inheritance. Abstracting away from the many still unknown details of how information is propagated through a plant during direct and indirect induction (Morrell and Kessler 2014) and using the proposed epigenetic inheritance-based acquisition schematic, we can roughly represent the anticipatory process leading up to the expression of herbivore resistance-mediating phenotypes as follows: First, a plant detects VOCs, S_1 , at its membrane; this results both in the induction of product S^* and the epigenetic modification of associated gene S^* . The continued presence of the VOC (or detecting it at a later time) amplifies the level of S^* beyond that which is required for autocatalysis. After a sufficiently high level of product S^* is reached, it affects gene S^{**} by amplifying the signal associated with S_2 (herbivore salivatory chemicals) that is needed to begin and sustain production of a defense-related gene product. When herbivore salivatory chemicals are detected at the plant's membrane, a faster and stronger response can be initiated due to the fact that gene S^{**} in the context of the larger regulatory network has been proactively primed for delivering a defense-related response.¹⁶

¹⁶ For an additional support for epigenetic inheritance-based model acquisition see Latzel and Münzbergova's (2018) study on anticipatory behavior of clonal plant *Fragaria vesca* and its basis in epigenetic transgenerational memory.

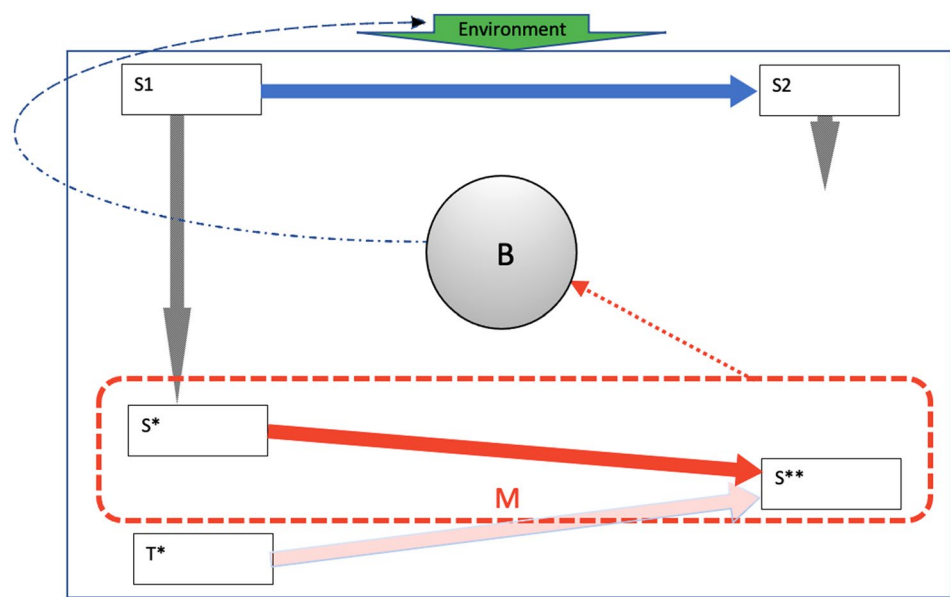
Let us now briefly turn to consider how associative learning results in the acquisition of anticipatory models within the lifetime of an organism.

Associative Learning-Based Acquisition

Associative learning may be characterized as “learning about the relationship between two separate stimuli, where the stimuli might range from concrete objects and events to abstract concepts, such as time, location, context, or categories” (Lafontaine et al. 2020). In keeping with the notion of correlational structure introduced in the second section and used above in what follows, I shall focus exclusively upon the relationship between event types. The two primary experimental procedures with which associative learning is studied are *Pavlovian conditioning* and *operant conditioning*. In Pavlovian conditioning a contingency relation is learned between a conditioned stimulus and an unconditioned stimulus. A conditioned stimulus is both valence neutral (e.g., auditory tone) and does not elicit a regular response. An unconditioned stimulus, on the other hand, is one that has some biological relevance and hence positive or negative valence for a given organism (e.g., food or electric shock) and will typically elicit an unconditioned response (e.g., salivation in the presence of food or aversive jumping behavior to shock). Pavlovian conditioning occurs when an organism is regularly exposed to the conditioned followed by unconditioned stimulus, such that via some internal associative linking of the two stimuli the unconditioned response—also called the conditioned response—can be elicited by the presence of the conditioned stimulus alone (e.g., an auditory tone triggers salivation in the absence of food) (Ginsburg and Jablonka 2009).

In contrast to the stimulus-stimulus contingency relation of Pavlovian conditioning, in operant conditioning a stimulus-response relation is learnt (e.g., food will follow from pushing a lever). After regularly being exposed to the stimulus-response pair and hence learning the association between the two, the learner engages in or avoids a particular response behavior (an operant) in expectation that the stimulus (a reinforcement or a punishment) will result from that behavior (e.g., push the lever for food) (Domjan 2018). In both of these forms of associative learning, it is in virtue of the fact that one event is a predictor of another that an organism's behavior is proactive. In the case of Pavlovian conditioning, it is the conditioned stimulus that regularly follows an unconditioned stimulus that a response is directed towards; in the case of operant conditioning, it is the yet-to-be encountered reinforcement (or punishment) regularly brought about by the learner's own behavior which that behavior is directed towards. As such, anticipation is part and parcel of successful associative learning.

Fig. 4 Schematic of associative learning-based anticipatory model acquisition: S1 and S2=regularly successive sensory states; S*, T*, and S**= internal states; M=the model; B=behavior. Prior to learning, the connections between T* and S** and S* and S** may be similar in strength. During learning, S* and S** become strongly linked due to the organism regularly experiencing S2 following S1, while the linkage between T* and S** decreases (or is lost altogether) due to the lack of encountering S2 following T1. S* and S** jointly instantiate M that elicits B prior to the occurrence of S2



With this in mind, associative learning-based acquisition can be characterized as follows: in learning stimulus-stimulus or stimulus-response contingencies, an organism acquires an anticipatory model of the environmental correlational structure that it regularly encounters in its lifetime. In other words, the process of one event *becoming linked to another* is analogous to acquiring an anticipatory model via associative learning.

Redescribing this form of model acquisition schematically and limiting the focus to Pavlovian conditioning: a sensory state, S2, brought about by an unconditioned stimulus, causes internal state S**, which encodes information about the presence of the unconditioned stimulus. Due to the biological relevance of the unconditioned stimulus, S** brings about an unconditioned behavioral response, B. Subsequently, the organism regularly encounters a conditioned stimulus that is registered by sensory state S1 and followed by S2. S1 gives rise to a distinct internal state, S*, that encodes information about the presence of the conditional stimulus. Given that S* is regularly followed by S2's causing S**, the two internal states S* and S** are linked. This linking between S* and S** is the instantiation of a *correlational rule* in the model which (statistically) captures the environmental structure of the events causing S1 and S2. That a model has been acquired is evidenced by the fact that S1 via the model can activate S** and hence elicit B prior to the occurrence of S2 (or even in the absence of S2) (see Fig. 4).

In comparison to mutation-based and epigenetic inheritance-based acquisition, one distinct feature of associative learning-based model acquisition is that such model acquisition brings with it the ability for individuals to plastically adapt to changing environmental structures, the correlational

(re)organization of which happens during an individual's lifetime. In other words, because associative learning occurs across ontogenetic timescales, it allows organisms to harness and extrapolate within a model possibly many, if not in some cases innumerable (cf. Ginsburg and Jablonka 2019), regular changes to correlational structure. Associative learning-based model acquisition allows for the rapid update of anticipatory models which takes the form of modified correlational rules and hence the pruning and modification of molecular pathways.¹⁷ This of course says a lot about the timescales at which the correlational structures that associative learning-based models capture occur; such structures are themselves occurring repeatedly across an organism's lifetime. Moreover, the fact that neural pathways operate on a millisecond-to-second timescale as compared to the minute-to-day timescale of gene regulatory networks suggests that the timescales of the correlational structures that neural

¹⁷ In neuronal organisms, this can be understood in terms of the strengthening of synaptic connections between neurons that are regularly and contiguously activated, thereby increasing the likelihood that when one neuron fires, the other will subsequently fire. The strength of other synaptic connections that may have been in place, but which have not been regularly activated are down-weighted (i.e., "pruned") by experience. This process of Hebbian learning is often expressed as neurons that "fire together, wire together." Given the evidence that associative learning is not limited to neuronal organisms, this notion of synaptic plasticity cannot be used to account for the underlying mechanisms of associative learning in non-neuronal organisms. I will assume that in the case of non-neuronal organisms, molecular activity plays the role of memory which predisposes metabolic pathways and regulatory networks to respond in ways that are history dependent, in much the same way that synaptic connection activity does in neuronal systems (cf. De la Fuente 2015; Gershman et al. 2021).

pathways can capture are significantly faster than those that are able to be captured by the molecular activity of gene regulatory networks. This difference is particularly relevant when considering the possibility that non-neuronal organisms engage in associative learning-based model acquisition and thus thinking about the timescales of the correlational structures that their internal molecular dynamics encode.

In associative learning-based acquisition it is the regularity between correlated environmental events (e.g., stimulus-stimulus or stimulus-response pairs) within the lifetime of an individual that determines whether or not (and how) an anticipatory model is acquired. In order for learning to occur, a correlational structure *may* have to be encountered by an individual many times. In contrast to both mutation-based and epigenetic inheritance-based acquisition, environmental regularity encountered across generations does not play a role in associative learning-based acquisition. As we shall see in the next subsection, however, Baldwin effect-acquisition represents an interesting case where an individual regularly encountering some correlational structure (within its lifetime) may lead to regular structure across generations, having an influence upon whether a model is acquired in later generations.

There are many examples of associative learning in the literature and given the straightforward relationship between such learning and associative learning-based anticipatory model acquisition, each example of the former may be used to exemplify a case in which the latter has occurred. As such, I will refrain from rehearsing any particular example(s) here and refer the interested reader to Ginsburg and Jablonka (2019). That said, it is interesting to note that as of today, examples of associative-like learning in non-neuronal organisms are rare and limited to a few studies on paramecia (Gelber 1952, 1958; Armus et al. 2006), amoebae (De la Fuente et al. 2019; Carrasco-Pujante et al. 2021), and a *possible* example in pea plants (Gagliano et al. 2016). This might be taken to suggest that anticipatory learning-based model acquisition is *largely* restricted to neuronal organisms. I would caution, however, that more ecologically valid testing conditions of such capacities in non-neuronal organisms are needed before pronouncing associative learning (and hence associative learning-based acquisition) in non-neuronal organisms to be something that is actually rare in nature. For the claim that associative learning is rarely found in non-neuronal organisms is certainly not equivalent to the claim that associative learning in non-neuronal organisms is rarely found in laboratory conditions that bear little (if any) resemblance at all to the natural habitats in which such learning might be homeostatically motivated.

Let us now move on to the final manner of anticipatory model acquisition that I shall consider in this article. In the next subsection, I will start by providing a general overview of the Baldwin effect and then go on show how this

mechanism might form the basis of one additional manner of acquiring anticipatory models.

The Baldwin Effect-Based Acquisition

The Baldwin effect refers to a hypothetical evolutionary process proposed by psychologist James Mark Baldwin (1896, 1902).¹⁸ It was introduced as a manner of supplementing Darwinian evolution by natural selection largely due to Baldwin's contention that random mutation by itself was unlikely to account for how all the variations that are selected for in nature are produced (Crispo 2007). The Baldwin effect, as such, posits a case in which some novel, nonheritable phenotype that is plastically induced in a subpopulation not only results in the increased survival and fitness of those individuals that exhibit it, but over many subsequent generations of selection acting in the direction of the phenotype, it becomes a heritable trait. The Baldwin effect may be construed in terms of a process that progresses through the three related stages: phenotypic accommodation, organic selection, and orthoplasy. The first stage, *phenotypic accommodation*, refers to the idea that a novel or unusual environmental input (challenge) induces a particular nonheritable phenotypic plastic response without genetic change. The second stage, *organic selection*, refers to the notion that the presence of an adaptive plastic phenotypic response in its inducing environment increases the survival of the organisms that are able to exploit it and those organisms as a result produce more offspring. Importantly, any developmental systems that converge on an adaptive phenotype more easily, earlier, and/or across a wider range of contexts, may be subject to organic selection and this may occur without necessarily seeing any significant change in gene frequencies across the population (cf. Nijhout et al. 2021). The third and last stage, *orthoplasy*, refers to the idea that over time, as a result of evolution by natural selection, the plastic response that was originally environmentally induced is selected for and genetically embedded such that the response becomes heritable. In other words, natural selection acts on "variations in the direction of plasticity" (Baldwin 1902, p. 37).

The Baldwin effect applies to a wide range of phenotypic accommodations that include learning and behavioral changes in addition to morphological and developmental changes.¹⁹ Moreover, insofar as the possibility that some of these plastic responses are epigenetic variations, there is potential for the Baldwin effect to also facilitate the genetic

¹⁸ The history of the Baldwin effect is not a straightforward matter. For the details see Depew (2000).

¹⁹ For striking empirical evidence regarding the Baldwinization of novel developmental and morphological accommodations in the North American house finch see Badyaev (2009).

fixation of epigenic responses (cf. Bateson and Gluckman 2011; see also Badyaev 2009).

With respect to learning, individual learned behaviors are subject to orthoplasia by way of natural selection acting on interacting metabolic, physiological, morphological, developmental, and/or behavioral recurrent trait complexes (West-Eberhard 2003), each element of which contributes to a learned behavior's coming about. Given that such a behavior increases survival and/or fitness, natural selection will favor any genetic differences with respect to the underlying trait complex that enable learning of the behavior to come about more easily (i.e., with less practice and/or earlier in development). Over time such selection will result in the fortification of the recurrent complex and the need for less and less learning in order for the behavior to be expressed in the population. The Baldwin effect thus posits a path from more efficient learning of a behavior in ancestral populations to heritable behavior that requires no initial learning on the part of that population's *distant* progeny. It should be emphasized, however, that the selection and genetic embedding of a plastic phenotype *does not entail that plasticity as a general feature is lost* (Bateson 2014). On the contrary, selection can change the mean trait values of a phenotype leaving the level of plasticity in the population untouched (Crispo 2007) or increase plasticity in the population given that the most plastic individuals are those that have the most extreme phenotypes and hence are the ones that are selected for (Baldwin 1902, pp. 36–37).²⁰ So although it might be the case that a specific phenotype after genetic encoding (i.e., genetic accommodation) becomes nonplastic, plasticity in a population can increase via its being directly subject to selection (Schlichting and Pigliucci 1993).

With a general understanding of the Baldwin effect to hand, we may now articulate *Baldwin effect-based acquisition*. This particular path to anticipatory model acquisition refers to the idea that a model that is initially acquired in virtue of plasticity (i.e., associative learning) during an individual's lifetime may be genetically encoded over time, resulting in a heritable anticipatory model. Using the three stages of the Baldwin effect to sharpen this more general description, Baldwin effect-based acquisition may be characterized as follows: during their lifetime, certain individuals in a population learn (via associative learning) a particular correlational relation between a regularly occurring environmental cue and outcome states and in doing so gain the

ability to anticipatorily behave in response to that cue. This is the phenotypic accommodation stage. Those individuals in the population that both learn this correlation more easily and/or earlier, and can respond adaptively with the appropriate AB across a wider range of contexts, survive longer and reproduce more than those that do not. This is the organic selection stage. Over generations, any genetic variation that makes learning the specific association and the respective AB easier is selected for, resulting in the eventual genetic encoding of the correlational structure and respective adaptive AB across the population.²¹ At this point, an anticipatory response to the cue is no longer dependent upon the individuals in the population establishing contingency links (stimulus-stimulus or stimulus-response) via associative learning. This is the orthoplasia stage.

One question that arises is whether Baldwin effect-based acquisition—as described—presupposes that an anticipatory model has already been acquired via associative learning (during the phenotypic accommodation stage). At first blush, such a presupposition would seem to be problematic given that it suggests that this form of model acquisition presupposes the very thing that it is meant to explain. This explanatory circularity, however, can be shown to merely be apparent by recognizing that although Baldwin effect-based acquisition includes associative learning (and hence another form of model acquisition), learning by itself is not sufficient for Baldwin effect-based acquisition to arise. The type of explanation that is being offered by Baldwin effect-based acquisition is directed at cases in which an organism exhibits AB, but that AB is not the direct result of learning on the part of *that* organism; rather, such behavior is the result of a history of (accelerated) learning on the part of that organism's ancestors in addition to their being subject to organic selection and orthoplasia. In contrast to this, the AB exhibited by members of the initial ancestral subpopulation of associative learners is not subject to explanation by Baldwin effect acquisition.

Crucially, what is thrown into relief by Baldwin effect-based acquisition is a possible path from an initial associative learning-based acquisition in individuals to a form of mutation-based acquisition across generations, the result of which is AB in individuals without their needing to engage in associative learning. As such, Baldwin effect-based acquisition may be represented by a two-part schematic. For example, we might start out with an individual organism that after moving into a new environment regularly encounters

²⁰ This somewhat deflates various arguments against the Baldwin effect that are based on the idea that as a result of the effect populations become “dumb” or inflexible (cf. Godfrey-Smith 2003). As remarked by Crispo (2007), it is phenotypic assimilation (Waddington 1953, 1957), an evolutionary mechanism that is often confused with the Baldwin effect, that entails a loss of plasticity via canalization.

²¹ One contender that may act as a possible steppingstone between learning a particular association and encoding that association in the activities of gene regulatory networks may be epigenetic inheritance mechanisms. For an overview as to how epigenetic modification may facilitate genetic fixation of particular plastic phenotypes more generally see Bateson and Gluckman (2011).

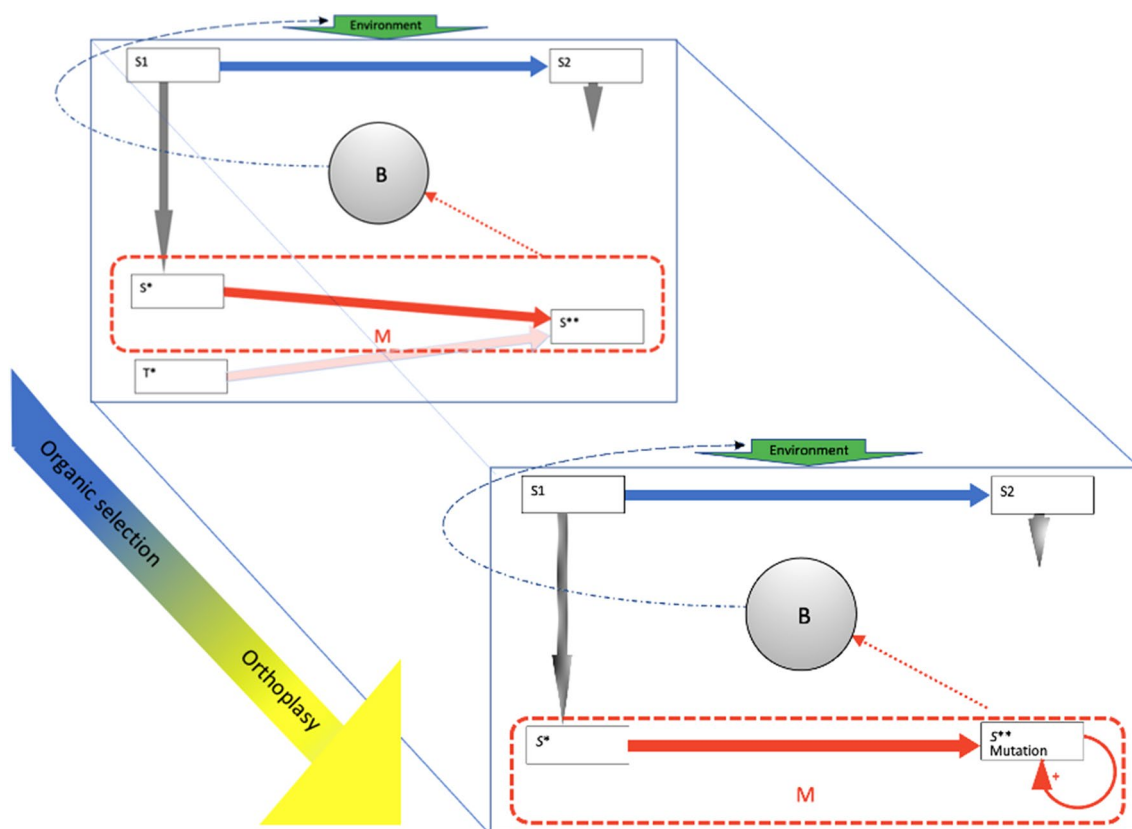


Fig. 5 Schematic of Baldwin effect-based anticipatory model acquisition: in the two stages of this form of anticipatory model acquisition (associative learning-based acquisition and mutation-based acquisition), all symbols have the same referents as they did in each respec-

tive form of model acquisition as described above. Here, M, firstly acquired via associative learning, is then by way of organic selection and subsequent orthoplasy embedded in the genome over multiple generations

some environmental event causing sensory state S1 (a conditioned stimulus) followed by another event causing sensory state S2 (an unconditioned stimulus). Given that S2 is a state that challenges the organism's sustained homeostatic maintenance, S1 elicits an anticipatory avoidance behavior in the organism prior to its encountering the negative unconditioned stimulus. This anticipatory avoidance behavior occurs in virtue of the fact that an internal state S**, one that is typically caused by S2, becomes linked to another internal state S* that is caused by S1. The linking of the two internal states is the result of the repeated activity of S** following S* such that a correlational rule is established via dynamics of the interacting states.

Up to this point, I have just described associative learning-based acquisition—a specific form of phenotypic accommodation. To this, we may imagine a subpopulation of individuals in the new environment that are similar to this organism in that they are able to acquire an anticipatory model via associative learning. Importantly, however,

some do it more easily and earlier in development than others. Given that experiencing S2 may be both common in the new environment and that it leads to deleterious effects, those organisms that do not acquire the anticipatory model or do so less efficiently will neither live as long nor reproduce as much as those organisms that acquire the model rapidly. After multiple generations in which this process of organic selection proceeds to select for better learners in the same environmental context, orthoplasy takes place; namely, any random changes in DNA sequences that might contribute to the learnt model becoming heritable will be selected for and effectively push the model down into the genes and into the activity of complexes of gene regulatory networks. As a result of this process occurring over multiple generations, a descendant organism when sensing S1 can respond with the appropriate avoidance behavior prior to encountering S2 via an inherited anticipatory model that is largely instantiated in

the very organization of its gene regulatory networks (S^* and S^{**}) (see Fig. 5).²²

Although there are no examples that I am aware of which *directly* support Baldwin effect-based acquisition, one study by Mery and Kaweki (2002) is of particular interest for evaluating such acquisition not only because it provides an example of the Baldwin effect (Hayes et al. 2020), but because it provides an opportunity to consider how a similar experimental design might be used to uncover Baldwin effect-based model acquisition. Deploying an experimental evolution design, Mery and Kaweki (2002) investigated the conditions in which an improved learning ability might evolve in the lesser fruit fly (*Drosophila melanogaster*). In this experiment, a group of 150 flies were first allowed the choice to lay eggs on either an orange medium with a small percentage of quinine (an aversive gustatory chemical) or on a pineapple medium without quinine. This first period allowed flies to learn an association between a flavor/odor (e.g., orange) and quinine, one that is relevant for the females' substrate choice for ovipositing; eggs are typically laid on the same substrates that are fed on by adult flies and thus feeding provides a chance to sample and evaluate the quality of the oviposition substrate.

Subsequently, during two additional consecutive three-hour periods these same 150 flies were given a choice between two fresh media (orange and pineapple), neither of which contained quinine. The proportion of eggs laid on each type of media in periods two and three was used to track oviposition preference. After the third period, 250 eggs were collected (i.e., selected) exclusively from the medium that was not adulterated with quinine in the first period. These three periods together were considered equivalent to one *Drosophila* generation. After incubating and hatching the eggs from the first (odd) generation, the new (even) generation was exposed to the same conditions over three periods, except that in the first period the quinine was paired with the pineapple medium and not the orange. The

same selection regime was applied. After 45 generations a comparison between a control population (i.e., flies that over the course of 45 generations were exposed to similar conditions but without quinine) and the experimental population was made in which both were presented with a choice between a quinine-free and a quinine-adulterated medium. This comparison revealed that the “experimental fly populations evolved an improved ability to associate the taste or smell of an oviposition medium with an aversive chemical cue (quinine) and to avoid this medium several hours later, when the cue was no longer present” (Mery and Kaweki 2002, p. 14,278).

These results provide strong evidence for Baldwinized learning (Heyes et al. 2020). *Drosophila*'s exposure to new environmental conditions is met by phenotypic accommodation (associative learning) that, via experimentally imposed selection, is pushed down into the genes over multiple generations. It is important to note, however, that it is improved learning (e.g., less time required for conditioning) itself and not the specific stimulus-(aversive)stimulus learnt pairs that is shown to be subject to the Baldwin effect, seeing as pair types alternate between even and odd generations over the course of the evolutionary experiment. With these details in mind, we are now in a position to formulate an important question: if the learning progressively improves when the *general* learning conditions (i.e., that selection is based on a quinine pairing) are reliable across generations, what might we expect when *specific* learning conditions such as orange-quinine and pineapple-no quinine are stable across generations? One hypothesis that calls out for testing—and one that is central to evaluating Baldwin effect-based acquisition—is this: not only might the comparative time required for conditioning in the experimental lineage be less than that of the control lineage, but we might also expect that a learned association between specific stimuli (orange-quinine) becomes genetically embedded over many generations under the selection regime constraints of experimental evolution. This orthoplasmy would be evidenced by the observation that the experimental lineages would increasingly opt for non-orange ovipositing substrates *without the need to be conditioned in the first period*. In other words, at some point, flies from the experimental lineages might engage in *non-learned anticipatory avoidance behavior*, flying directly to any other available medium to feed and lay their eggs when given a choice between it and an orange medium. Such results would support the notion of Baldwin effect-based model acquisition.

Taking stock: in this section I have presented four different yet interrelated paths to AB, each of these paths representing different ways that organisms might acquire anticipatory models (see Table 1). In laying these different paths out for inspection on the same theoretical table, it should be evident that although they occur across a range

²² With respect to the qualification “largely”: strictly speaking, in neuronal organisms the learned association (and associated AB) is never fully “handed over” to the genes; rather, the genome plays an increasingly central role in stabilizing the association and the respective AB. This stabilization may take the form of heritable genomic changes or epigenetic control exerted on the genome that makes the at-one-time experience-dependent weighting of certain synaptic pathways less and less dependent upon learning. For example, one way that gene activity may affect weighting of particular pathways or pathway structures is by increasing the amount of metabolic energy directed at and governing particular synaptic connections (see Yuan et al. 2018). Additionally, over many generations developmental processes (e.g., hormone signalling) triggered by gene regulatory networks may provide a manner for regularly encountered environmental structure to influence synaptic pathway structure and function (Sinha et al. 2020), such that the development of those synaptic structures becomes part of the regular developmental trajectory of a population (as long as the AB such structure engenders remains adaptive).

Table 1 Different ways to acquire anticipatory models

Mutation-based acquisition	Random genetic mutations result in a regulatory network's being able to harness environmental correlational structure, allowing for anticipatory behavioral phenotypes that can be transmitted across generations
Epigenetic inheritance-based acquisition	Heritable epigenetic variation via epigenetic mechanisms as situated within a gene regulatory network can harness environmental correlational structure and thus lead to adaptive anticipatory behavioral phenotypes that can be transmitted across generations
Associative learning-based acquisition	Learning stimulus-stimulus or stimulus-response contingencies on the part of an organism within the course of its lifetime results in non-heritable anticipatory behavioral phenotypes
Baldwin effect-based acquisition	A non-heritable anticipatory behavioral phenotype acquired via associative learning-based acquisition is embedded genetically over multiple generations due to its adaptive benefits, resulting in a heritable anticipatory model and a heritable anticipatory behavioral phenotype

of timescales (from ontogeny to phylogeny), these ways of acquiring anticipatory models often and regularly interact, constraining and/or enabling one another.

Given these characterizations of four paths to anticipatory behavior, we may now reason to some select constraints that may serve as guidance for the identification of each form of model acquisition in nature. This is the task of the next and final section.

Guidance for Model Acquisition Identification

Assuming that the characterizations of the four manners of acquiring anticipatory models are accurate, they provide insight into the particular conditions under which each of them might be expected to be found in nature. Understanding these conditions is particularly important for identifying and investigating those forms of model acquisition that occur over phylogenetic timescales: such conditions can be used to inform the design of experimental evolutionary studies, offering substantial empirical support to theorizing about such forms of model acquisition and the historical environments in which may have arisen.

With respect to mutation-based model acquisition, we might expect it take place in organisms that occupy environments in which (survival-relevant) correlational structure has been stable over multiple generations. If, for example, some environmental correlational structure only lasts half the average lifetime of an individual organism and never shows up again for that individual or its progeny, then this would not be the kind of correlation that mutation-based model acquisition would be able to capture. We might also expect mutation-based acquisition to occur in environments where the trajectory between two correlated events (as detected by S1 to S2) is slower than the relevant phenotype-modifying gene expression/repression that the cue event triggers.

Moving on to epigenetic inheritance-based model acquisition: like mutation-based acquisition, this form of

acquisition might be expected to take place in organisms that occupy environments in which correlational structure has been stable over multiple generations. However, models acquired in this manner can be transmitted even in the absence of an organism's encountering the correlational structure that the model captures and hence in the absence of needing to actualize model-based AB. For example, a plant that fails to experience herbivory may still instantiate an anticipatory model that was acquired from its progenitor, a mother plant that acquired a model in virtue of its encounters with herbivores. Relatedly, epigenetic inheritance-based model acquisition might be expected in those organisms that are *regularly* subject to the same environmental stressors that their progenitors also face/faced. Like mutation-based acquisition, we would also expect epigenetic inheritance-based model acquisition in environments where the trajectory between correlated events is slower than the phenotype-modifying gene expression/repression that compensates for the yet-to-be encountered event.

With respect to associative learning-based acquisition, we might expect to find this kind of acquisition independently of whether or not progeny encounter the same correlational environmental structure as their progenitors; this is because associative learning-based acquisition does not involve model transmission across generations. We also might expect to find such model acquisition to occur in environments where the trajectory across correlated events is slower than the gene expression/repression that underwrites phenotypic plasticity and/or slower than the induced motor response to the cue. To see why this is the case, consider the following: an auditory tone and an aversive shock are paired such that the onset of both stimuli is always identical over the course of the training period. In this case, although the tone might become associated with the shock (or the shock associated with the tone), since there is no time gap between the onsets of the two stimuli and hence no possibility for the tone to act as a cue for the shock, there would be no experiential context for an anticipatory response and hence no experiential context for an anticipatory model to be acquired.

Lastly, we would expect Baldwin effect-based acquisition to be found only in those organisms that have the capacity to engage in associative learning. Akin to mutation-based and epigenetic inheritance-based acquisition, Baldwin effect-based acquisition might be expected to occur in organisms that occupy environments with a stability of correlational structure that spans multiple generations. Particular to Baldwin effect-based acquisition is that this form of acquisition might be expected to arise in populations of organisms that have variable learning rates. Were there no variance or outliers in the population learning-rate distribution, then organic selection would not get off the ground. This is not to suggest that there are cases in which no variance in learning rate across a population exists; rather it is to suggest that Baldwin effect-based acquisition might be more likely to occur in populations where the variable learning rate is more extreme, creating an adaptive difference that can be selected for.

Conclusion

Under the assumption that the AB found in biological systems is underwritten by anticipatory models, the fact that AB is found across the biological board suggests that anticipatory models may be a ubiquitous feature of life. The aim of this article has been to explore four non-exhaustive ways that anticipatory models might be acquired in a range of living systems, and by doing so provide a sharper picture of AB and the conditions under which it arises across different timescales. By articulating some of the various ways that anticipatory models might be acquired, the learning versus evolution discussion is brought forward by throwing light upon interaction between ontogenetic and phylogenetic adaptive processes; learning, although happening exclusively over ontogenetic timescales, is contextually embedded in and influenced by the results of acquisition that has occurred at evolutionary timescales; model acquisition occurring at phylogenetic timescales can be influenced by plasticity occurring within the lifetime of an individual. This kind of bidirectional causal influence is often overlooked and yet should hardly be surprising given that an individual's phenotype is neither wholly the product of its interactions nor wholly the product of the interactions of its progenitors. Although the idea that phenotypical phenomena come about via the interaction of a host of causal influences across different levels (from genetic to environmental) is not new (cf. Waddington 1957; Goodwin 1994; Oyama 2000), its importance for the project of understanding the intricacies of biological adaptive processes cannot be understated. The many paths to anticipatory behavior are no exception.

Acknowledgments I would like to thank the following people for their invaluable feedback on earlier drafts of this article and for all of the discussions that helped to refine many of the ideas presented here: Pamela Lyon, Eva Jablonka, Michael Levin, Giovanni Pezzulo, Carrie Figdor, Nick Brancazio, HG Döbereiner, Jan Baedke, Tobias Schlicht, Maxim Raginsky, and Benjamin Little. I would also like to thank two anonymous reviewers for their detailed comments and very helpful suggestions.

Author Contributions The author confirms being the sole contributor of this work.

Funding Open Access funding enabled and organized by Projekt DEAL. This research has been supported by the Alexander Von Humboldt Foundation.

Declarations

Conflict of interest The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Anastasiadi D, Piferrer F (2019) Epimutations in developmental genes underlie the onset of domestication in farmed European sea bass. *Mol Biol Evol* 36(10):2252–2264. <https://doi.org/10.1093/molbev/msz153>
- Arbib MA (1992) Schema theory. In: Shapiro S (ed) *Encyclopedia of artificial intelligence*, vol 2, 2nd edn. Wiley, Chichester, pp 1427–1443
- Armus HL, Montgomery AR, Gurney RL (2006) Discrimination learning and extinction in paramecia (*P. caudatum*). *Psychol Rep* 98(3):705–711. <https://doi.org/10.2466/pr0.98.3.705-711>
- Badyaev AV (2009) Evolutionary significance of phenotypic accommodation in novel environments: an empirical test of the Baldwin effect. *Philos Trans R Soc Lond B Biol Sci* 364(1520):1125–1141. <https://doi.org/10.1098/rstb.2008.0285>
- Baldwin JM (1896) A new factor in evolution. *Am Nat* 30:441–451. doi:<https://doi.org/10.1086/276408>
- Baldwin JM (1902) *Development and evolution*. Macmillan, New York
- Bartlett FC (1932) *Remembering: a study in experimental and social psychology*. Cambridge University Press, Cambridge
- Bateson P, Gluckman P (2011) *Plasticity, robustness, development and evolution*. Cambridge University Press, Cambridge
- Bateson P (2014) New thinking about biological evolution. *Biol J Linn Soc* 112:268–275

- Bechtel W (2011) Representing time of day in circadian clocks. In: Newen A, Bartels A, Jung E-M (eds) Knowledge and representation. CSLI Publications, Palo Alto, pp 129–162
- Bernhardt JR, O'Connor MI, Sunday JM, Gonzalez A (2020) Life in fluctuating environments. *Phil Trans R Soc B* 375:20190454. <https://doi.org/10.1098/rstb.2019.0454>
- Bickhard M (2009) The interactivist model. *Synthese* 166:547–591. <https://doi.org/10.1007/s11229-008-9375>
- Brunner AM, Nanni P, Mansuy IM (2014) Epigenetic marking of sperm by post-translational modification of histones and protamines. *Epigenetics & Chromatin* 7:2. <https://doi.org/10.1186/1756-8935-7-2>
- Cannon WB (1932) The wisdom of the body. Norton, New York
- Cantone I, Fisher AG (2013) Epigenetic programming and reprogramming during development. *Nat Struct Mol Biol* 20:282–289. <https://doi.org/10.1038/nsmb.2489>
- Carrasco-Pujante J, Bringas C, Malaina I, Fedetz M, Martínez L, Pérez-Yarza G et al (2021) Associative conditioning is a robust systemic behavior in unicellular organisms: an interspecies comparison. *Front Microbiol* 12:707086. <https://doi.org/10.3389/fmicb.2021.707086>
- Castel SE, Martienssen RA (2013) RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nat Rev Genet* 14:100–112
- Castelfranchi C (2005) Mind as an anticipatory device: for a theory of expectations. In: De Gregorio M, Di Maio V, Frucci M, Musio C (eds) BVAI 2005. LNCS, vol 3704. Springer, Heidelberg, pp 258–276
- Cedar H, Bergman Y (2009) Linking DNA methylation and histone modification: patterns and paradigms. *Nat Rev Genet* 10:295–234
- Champagne FA, Meaney MJ (2006) Stress during gestation alters postpartum maternal care and the development of the offspring in a rodent model. *Biol Psychiatry* 59:1227–1235
- Chandler VL (2007) Paramutation: from maize to mice. *Cell* 128(4):641–645
- Clark A (2016) Surfing uncertainty: prediction, action, and the embodied mind. Oxford University Press, Oxford
- Corcoran AW, Pezzulo G, Hohwy J (2020) From allostatic agents to counterfactual cognisers: active inference, biological regulation, and the origins of cognition. *Biology and Philosophy* 35:32. <https://doi.org/10.1007/s10539-020-09746-2>
- Craik KJW (1943) The nature of explanation. Cambridge University Press, Cambridge
- Crick F (1984) Neurobiology: memory and molecular turnover. *Nature* 312:101. DOI: <https://doi.org/10.1038/312101a0>
- Crispo E (2007) The Baldwin effect and genetic assimilation: revisiting two mechanisms of evolutionary change mediated by phenotypic plasticity. *Evolution* 61(11):2469–2479. <https://doi.org/10.1111/j.1558-5646.2007.00203.x>
- De la Fuente IM (2015) Elements of the cellular metabolic structure. *Front Mol Biosci* 2:16. doi: <https://doi.org/10.3389/fmolb.2015.00016>
- De la Fuente IM, Bringas C, Malaina I, Fedetz M, Carrasco-Pujante J, Morales M et al (2019) Evidence of conditioned behavior in amoebae. *Nat Commun* 10:3690. <https://doi.org/10.1038/s41467-019-11677-w>
- Deacon T (2012) Incomplete nature: how mind emerged from matter, 1st edn. Norton, New York
- Deans C (2021) Biological prescience: the role of anticipation in organismal processes. *Front Physiol* 12:672457. <https://doi.org/10.3389/fphys.2021.672457>
- Depew DJ (2003) Baldwin and his many effects. In: Weber BH, Depew DJ (eds) Evolution and learning: the Baldwin Effect reconsidered. MIT Press, Cambridge, pp 3–31
- Dhar R, Sägerser R, Weikert C, Wagner A (2013) Yeast adapts to a changing stressful environment by evolving cross-protection and anticipatory gene regulation. *Mol Biol Evol* 30:573–588. [10.1093/molbev/mss253](https://doi.org/10.1093/molbev/mss253)
- Domjan M (2018) The essentials of conditioning and learning, 4th edn. American Psychological Association, Washington
- Donaldson-Matasci MC, Bergstrom CT, Lachmann M (2010) The fitness value of information. *Oikos* 119:219–230. doi: <https://doi.org/10.1111/j.1600-0706.2009.17781.x>
- Drescher GL (1991) Made-Up minds: a constructivist approach to artificial intelligence. MIT Press, Cambridge
- Dubnau D, Losick R (2006) Bistability in bacteria. *Mol Microbiol* 61(3):564–572
- Edgar RS, Green EW, Zhao Y, van Ooijen G, Olmedo M, Qin X et al (2012) Peroxiredoxins are conserved markers of circadian rhythms. *Nature* 485(7399):459–464. <https://doi.org/10.1038/nature11088>
- Fallet M, Luquet E, David P, Cosseau C (2020) Epigenetic inheritance and intergenerational effects in mollusks. *Gene* 729:144166. <https://doi.org/10.1016/j.gene.2019.144166>
- Freddolino PL, Tavazoie S (2012) Beyond homeostasis: a predictive-dynamic framework for understanding cellular behavior. *Annu Rev Cell Dev Biol* 28:363–384. doi: <https://doi.org/10.1146/annurev-cellbio-092910-154129>
- Friston KJ (2010) The free-energy principle: a unified brain theory? *Nat Rev Neurosci* 11:127–138
- Friston KJ (2019) A free energy principle for a particular physics. ArXiv: 1906.10184:1–148
- Friston KJ, FitzGerald T, Rigoli F, Schwartenbeck P, Pezzulo G (2017) Active inference: a process theory. *Neural Comput* 29(1):1–49. doi: https://doi.org/10.1162/NECO_a_00912
- Gagliano M, Vyazovskiy VV, Borbély AA, Grimonprez M, Depczynski M (2016) Learning by association in plants. *Sci Rep* 6:38427. <https://doi.org/10.1038/srep38427>
- Gallistel CR (2017) The coding question. *Trends Cognit Sci* 21:498–508. <https://doi.org/10.1016/j.tics.2017.04.012>
- Gelber B (1952) Investigations of the behavior of *Paramecium aurelia*. I. modification of behavior after training with reinforcement. *J Comp Physiol Psychol* 45:58–65. <https://doi.org/10.1037/h0063093>
- Gelber B (1958) Retention in *Paramecium aurelia*. *J Comp Physiol Psychol* 51:110–115. <https://doi.org/10.1037/h0049093>
- Gershman SJ, Balbi PEM, Gallistel CR, Gunawardena J (2021) Reconsidering the evidence for learning in single cells. *eLife* 10:e61907. <https://doi.org/10.7554/eLife.61907>
- Giaever G, Chu AM, Ni L, Connelly C, Riles L et al (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418:387–391
- Gilbert SF, Sarkar S (2000) Embracing complexity: organicism for the 21st century. *Dev Dyn* 219(1):1–9
- Ginsburg S, Jablonka E (2009) Epigenetic learning in non-neural organisms. *J Biosci* 34:633–646
- Ginsburg S, Jablonka E (2019) The evolution of the sensitive soul: learning and the origin of consciousness. MIT Press, Cambridge
- Godfrey-Smith P (2003) Between Baldwin skepticism and Baldwin boosterism. In: Weber BH, Depew DJ (eds) Evolution and learning: the Baldwin Effect reconsidered. MIT Press, Cambridge, pp 53–67
- Goodwin BC (1994) How the Leopard changed its spots: the evolution of complexity. Princeton University Press, Princeton
- Grewal SIS, Jia ST (2007) Heterochromatin revisited. *Nat Rev Genet* 8(1):35–46
- Harvey ZH, Chakravarty AK, Fuita RA, Jarosz DF (2020) A prion epigenetic switch establishes an active chromatin state. *Cell* 180:928–940
- Heard E, Martienssen RA (2014) Transgenerational epigenetic inheritance: myths and mechanisms. *Cell* 157(1):95–109

- Herman JJ, Spencer HG, Donohue K, Sultan SE (2013) How stable 'should' epigenetic modifications be? Insights from adaptive plasticity and bet hedging. *Evolution* 68:632–643. doi:<https://doi.org/10.1111/evo.12324>
- Heyes C, Chater N, Dwyer DM (2020) Sinking in: the peripheral Baldwinisation of human cognition. *Trends Cogn Sci* 24(11):1–18. <https://doi.org/10.1016/j.tics.2020.08.006>
- Holeski LM, Jander G, Agrawal AA (2012) Transgenerational defense induction and epigenetic inheritance in plants. *Trends Ecol Evol* 27:618–626
- Holliday R (1987) The inheritance of epigenetic defects. *Science* 238(4824):163–70. <https://doi.org/10.1126/science.3310230>
- Hu L, Xiao P, Jiang Y, Dong M, Chen Z, Li H et al (2018) Transgenerational epigenetic inheritance under environmental stress by genome-wide DNA methylation profiling in cyanobacterium. *Front Microbiol* 9:1479. <https://doi.org/10.3389/fmicb.2018.01479>
- Jablonka E (2017) The evolutionary implications of epigenetic inheritance. *Interface Focus* 7:20160135. <https://doi.org/10.1098/rsfs.2016.0135>
- Jablonka E, Lamb M (2005) *Evolution in four dimensions: genetic, epigenetic, behavioural, and symbolic variation in the history of life*. MIT Press, Cambridge
- Jablonka E, Lamb M (2020) *Inheritance systems and the extended evolutionary synthesis (elements in the philosophy of biology)*. Cambridge University Press, Cambridge
- Jablonka E, Raz G (2008) Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity. *Q Rev Biol* 84(2):131–176. <https://doi.org/10.1086/598822>
- Kauffman S (2000) *Investigations*. Oxford University Press, New York
- Kessler A, Halitschke R, Diezel K, Baldwin IT (2006) Priming of plant defense responses in nature by airborne signaling between *Artemisia tridentata* and *Nicotiana attenuata*. *Oecologia* 148(2):280–292. <https://doi.org/10.1007/s00442-006-0365-8>
- Kiefer A, Hohwy J (2018) Content and misrepresentation in hierarchical generative models. *Synthese* 195:2387–2415
- Kim J, Felton GW (2013) Priming of antiherbivore defensive responses in plants. *Insect Sci* 20(3):273–285. <https://doi.org/10.1111/j.1744-7917.2012.01584.x>
- Klosinska MM, Crutchfield CA, Bradley PH, Rabinowitz JD, Broach JR (2011) Yeast cells can access distinct quiescent states. *Genes Dev* 25:336–349
- Kozioł M, Rinn JL (2010) RNA traffic control of chromatin complexes. *Curr Opin Genet Dev* 20:142–148
- Lachmann M, Jablonka E (1996) The inheritance of phenotypes: an adaptation to fluctuating environments. *J Theor Biol* 181:1–9. doi:<https://doi.org/10.1006/jtbi.1996.0109>
- Lafontaine MP, Knoth IS, Lippé S (2020) Learning abilities. In: Gallagher Anne, Bulteau Christine, Cohen David, Michaud Jacques L (eds) *Handbook of clinical neurology*, vol 173. Elsevier, Amsterdam, pp 241–254
- Laforsch C, Beccara L, Tollrian R (2006) Inducible defences: the relevance of chemical alarm cues in *Daphnia*. *Limnol Oceanogr* 51:1466–1472
- Landmann S, Holmes CM, Tikhonov M (2021) A simple regulatory architecture allows learning the statistical structure of a changing environment. *eLife* 10:e67455. <https://doi.org/10.7554/eLife.67455>
- Latzel V, Münzbergová Z (2018) Anticipatory behavior of the clonal plant *Fragaria vesca*. *Front Plant Sci* 9:1847. <https://doi.org/10.3389/fpls.2018.01847>
- Levin M (2019) The computational boundary of a "self": developmental bioelectricity drives multicellularity and scale-free cognition. *Front Psychol* 10:2688. <https://doi.org/10.3389/fpsyg.2019.02688>
- Lim JP, Brunet A (2013) Bridging the transgenerational gap with epigenetic memory. *Trends Genet* 29:176–186. doi:<https://doi.org/10.1016/j.tig.2012.12.008>. PMID: 23410786
- Louie AH (2010) Robert Rosen's anticipatory systems. *Foresight* 12(3):18–29
- Louie AH (2012) Anticipation in (M,R)-systems. *Int J Gen Syst* 41(1):5–22
- Lyon P (2006) The biogenic approach to cognition. *Cogn Process* 7:11–29
- Markel K (2020) Lack of evidence for associative learning in pea plants. *eLife* 9:e57614. <https://doi.org/10.7554/eLife.57614>
- Maturana HR, Varela FJ (1980) *Autopoiesis and cognition: the realization of the living*. Reidel, Dordrecht
- Mery F, Kawecki TJ (2002) Experimental evolution of learning ability in fruit flies. *Proc Natl Acad Sci USA* 99:14274–14279
- Mitchell A, Lim W (2016) Cellular perception and misperception: internal models for decision-making shaped by evolutionary experience. *BioEssays* 38:845–849. <https://doi.org/10.1002/bies.201600090>
- Mitchell A, Romano GH, Groisman B, Yona A, Dekel E, Kupiec M, Dahan O, Pilpel Y (2009) Adaptive prediction of environmental changes by microorganisms. *Nature* 460(7252):220–224
- Monroe JG, Srikanth T, Carbonell-Bejerano P, Becker C, Lensink M, Exposito-Alonso M et al (2022) Mutation bias reflects natural selection in *Arabidopsis thaliana*. *Nature* 602:101–105. <https://doi.org/10.1038/s41586-021-04269-6>
- Morrell K, Kessler A (2014) The scent of danger: volatile-mediated information transfer and defence priming in plants. *Biochem (Land)* 36(5):26–31
- Nasuto SJ, Hayashi Y (2016) Anticipation: beyond synthetic biology and cognitive robotics. *Biosystems* 148:22–31. doi:<https://doi.org/10.1016/j.biosystems.2016.07.011>
- Neisser U (1976) *Cognition and reality: principles and implications of cognitive psychology*. Freeman, New York
- Nestler EJ (2016) Transgenerational epigenetic contributions to stress responses: fact or fiction? *PLoS Biol* 14(3):e1002426. doi:<https://doi.org/10.1371/journal.pbio.1002426>
- Nijhout HF, Kudla AM, Hazelwood CC (2021) Genetic assimilation and accommodation: models and mechanisms. *Curr Top Dev Biol* 141:337–369. <https://doi.org/10.1016/bs.ctdb.2020.11.006>
- Novoplansky A, Cohen D, Sachs T (1990) How portulaca seedlings avoid their neighbors. *Oecologia* 82:490–493
- Nowacki M, Vijayan V, Zhou Y, Schotanus K, Doak TG, Landweber LF (2008) RNA-mediated epigenetic programming of a genome-rearrangement pathway. *Nature* 451(7175):153–159
- Oyama S (2000) *The ontogeny of information: developmental systems and evolution*, 2nd edn. Duke University Press, Durham. <https://doi.org/10.1215/9780822380665>
- Pezzulo G, Rigoli F, Friston K (2015) Active inference, homeostatic regulation and adaptive behavioural control. *Prog Neurobiol* 134:17–35
- Pezzulo G (2008) Coordinating with the future: the anticipatory nature of representation. *Mind Mach* 18:179–220
- Piaget J (1970) *Genetic epistemology*. Columbia University Press, New York
- Poli R (2009) The many aspects of anticipation. *Foresight* 12(3):7–17. <https://doi.org/10.1108/14636681011049839>
- Poli R (2010) An introduction to the ontology of anticipation. *Futures* 42(7):769–776. <https://doi.org/10.1016/j.futures.2010.04.028>
- Ptashne M (2013) Epigenetics: core misconception. *Proc Natl Acad Sci USA* 110(18):7101–7103
- Radman-Livaja M, Liu CL, Friedman N, Schreiber SL, Rando OJ (2010) Replication and active demethylation represent partially overlapping mechanisms for erasure of H3K4me3 in budding yeast. *PLoS Genet* 6(2):e1000837

- Ramstead MJD, Kirchhoff MD, Friston K (2019) A tale of two densities: active inference is enactive inference. *Adapt Behav*. <https://doi.org/10.1177/1059712319862774>
- Rechavi O, Minevich G, Hobert O (2011) Transgenerational inheritance of an acquired small RNA-based antiviral response in *C. elegans*. *Cell* 147(6):1248–1256. <https://doi.org/10.1016/j.cell.2011.10.042>
- Riegler A (2001) The role of anticipation in cognition. In: Dubois DM (ed) *Computing anticipatory systems*. AIP Proceedings, Melville, pp 534–541
- Ringrose L, Paro R (2007) Polycomb/trithorax response elements and epigenetic memory of cell identity. *Development* 134(2):223–232
- Rivoire O, Leibler S (1949) 2014 A model for the generation and transmission of variation in evolution. *Proc Natl Acad Sci USA* 111:E1940–E1949. <https://doi.org/10.1073/pnas.1323901111>
- Roberts BT, Wickner RB (2003) Heritable activity: a prion that propagates by covalent autoactivation. *Genes and Development* 17(17):2083–2087
- Rodaki A, Bohovych IM, Enjalbert B, Young T, Odds FC, Gow NA, Brown AJ (2009) Glucose promotes stress resistance in the fungal pathogen *Candida albicans*. *Mol Biol Cell* 20(22):4845–4855. <https://doi.org/10.1091/mbc.e09-01-0002>
- Rodgers AB, Morgan CP, Bronson SL, Revello S, Bale TL (2013) Paternal stress exposure alters sperm microRNA content and reprograms offspring HPA stress axis regulation. *J Neurosci* 33:9003–9012. <https://doi.org/10.1523/JNEUROSCI.0914-13.2013>
- Rosen R (2002) The roles of necessity in biology. In: Casti J, Kariqvist A (eds) *Newton to Aristotle: towards a theory of models for living systems*. Birkhauser, Boston, pp 11–38
- Rosen R (1985/2012) *Anticipatory systems: philosophical, mathematical, and methodological foundations*. Pergamon, Oxford
- Rosen J (2009) Robert Rosen's anticipatory systems theory: the art and science of thinking ahead. *Proceedings of the 53rd Annual Meeting of the ISSS – 2009, Brisbane, Australia, 1(1)*. Retrieved from <https://www.journals.iss.org/index.php/proceedings53rd/article/view/1249>
- Saigusa, Tero A, Nakagki T, Kuramoto Y (2008) Amoebae anticipate periodic events. *Phys Rev Lett* 100:18101
- Schild S, Tamayo R, Nelson EJ, Qadri F, Calderwood SB, Camilli A (2007) Genes induced late in infection increase fitness of *Vibrio cholerae* after release into the environment. *Cell Host Microbe* 2(4):264–277. <https://doi.org/10.1016/j.chom.2007.09.004>
- Schlichting CD, Pigliucci M (1993) Control of phenotypic plasticity via regulatory genes. *Am Nat* 142(2):366–370. <https://doi.org/10.1086/285543>
- Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G (2007) Genome regulation by polycomb and trithorax proteins. *Cell* 128(4):735–745
- Sims M (2021) A continuum of intentionality: linking the biogenic and anthropogenic approaches to cognition. *Biol Philos* 36:51. doi: <https://doi.org/10.1007/s10539-021-09827-w>
- Sims M, Pezzulo G (2021) Modelling ourselves: what the free energy principle reveals about our implicit notions of representation. *Synthese* 199:7801–7833
- Sinha S, Jones BM, Traniello IM, Bukhari SA, Halfon MS, Hofmann HA et al (2020) Behavior-related gene regulatory networks: a new level of organization in the brain. *PNAS* 117:23270–23279
- Sobral M, Sampedro L, Neylan I, Siemens D, Dirzo R (2021) Phenotypic plasticity in plant defense across life stages: inducibility, transgenerational induction, and transgenerational priming in wild radish. *PNAS* 118(33):e2005865118
- Soen Y, Knafo M, Elgart M (2015) A principle of organization which facilitates broad Lamarckian-like adaptations by improvisation. *Biol Direct* 10:68
- Tagkopoulou I, Liu Y-C, Tavazoie S (2008) Predictive behavior within microbial genetic networks. *Science* 320(5881):1313–1317
- Tschantz A, Seth AK, Buckley CL (2020) Learning action-oriented models through active inference. *PLoS Comput Biol* 16(4):e1007805. <https://doi.org/10.1371/journal.pcbi.1007805>
- Van der Heijden GW, Derijck A, Ramos L, Giele M, Van der Vlag J, de Boer P (2006) Transmission of modified nucleosomes from the mouse male germline to the zygote and subsequent remodeling of paternal chromatin. *Dev Biol* 298(2):458–469
- Waddington CH (1953) Genetic assimilation of an acquired character. *Evolution* 7(2):118–126
- Waddington CH (1957) *The strategy of the genes*. Allen & Unwin, London
- Watson R, Szathmary S (2016) How can evolution learn? *Trends Ecol Evol* 31(2):147–157. doi: <https://doi.org/10.1016/j.tree.2015.11.009>
- West-Eberhard MJ (2003) *Developmental plasticity and evolution*. Oxford University Press, Oxford
- Yu A, Lepère G, Jay F, Wang J, Bapaume L, Wang Y et al (2013) Dynamics and biological relevance of DNA demethylation in Arabidopsis antibacterial defense. *Proc Natl Acad Sci USA* 110(6):2389–2394. <https://doi.org/10.1073/pnas.1211757110>
- Yuan Y, Huo H, Zhao P, Liu J, Liu J, Xing F, Fang T (2018) Constraints of metabolic energy on the number of synaptic connections of neurons and the density of neuronal networks. *Front Comput Neurosci* 12:91. <https://doi.org/10.3389/fncom.2018.00091>
- Zordan RE, Galgoczy DJ, Johnson AD (2006) Epigenetic properties of white-opaque switching in *Candida albicans* are based on a self-sustaining transcriptional feedback loop. *Proc Natl Acad Sci USA* 103(34):12807–12812

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.