

Beyond Diffusion: Bayesian Learning Strategies in Single-Cell Life

Maor Knafo  and Elena Casacuberta 

Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra), Pg Maritim Barceloneta 37, 08003 Barcelona, Spain

Iñaki Ruiz-Trillo 

ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain

and Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra), Pg Maritim Barceloneta 37, 08003 Barcelona, Spain



(Received 19 October 2025; accepted 24 April 2026; published 5 June 2026)

Evolutionary change is shaped not only by genetic variation and selection across generations but also by phenotypic plasticity, niche construction, and nongenetic inheritance. Here we develop a quantitative, falsifiable framework that models individual organisms as learners that infer hidden environmental states from noisy cues and update phenotype strategies within a lifetime. We introduce Markovian, Bayesian Agents (MBAs), which combine regret-gated stochastic exploration with genetic assimilation, and compare them to nonlearning Blind Agents (BAs). Agent-based simulations reveal that MBAs outperform BAs when cue-stress correlations are high, but lose their advantage above a critical environmental stochasticity threshold ($\varepsilon \approx 0.2$). We then test a central prediction in the unicellular holozoan *Capsaspora owczarzaki* using a smart-incubator conditioning assay: cultures trained with predictive cue-stress pairing show significantly reduced mortality relative to nonpredictive controls (Mann-Whitney U , $p = 1.07 \times 10^{-8}$). Together, model and experiment link within-lifetime inference to cross-generational adaptation in a single-celled system.

DOI: [10.1103/5p5z-t8qy](https://doi.org/10.1103/5p5z-t8qy)

I. INTRODUCTION

One of life's most remarkable features is its persistent and adaptive resilience in the face of constant environmental fluctuations [1–3]. Organisms across all biological domains routinely overcome severe internal and external stressors, highlighting fundamental adaptive mechanisms that allow rapid and flexible responses [4,5]. Understanding precisely how organisms accomplish this resilience remains a central question in biology, essential for unraveling the nature of adaptation itself.

Classical quantitative genetic models, particularly the canonical Wright-Fisher [6] and Kimura [7] frameworks, conceptualize adaptation as a gradual, population-level process driven by stochastic genetic mutations. These mutations, selected across generations, slowly refine organismal fitness in stable or nearly optimal environments, situations captured well by “Red Queen” dynamics [8–10], where continual adaptation is necessary to maintain relative fitness. Within these models, genotype and phenotype are tightly and deterministically linked, with each mutation mapping directly to a fixed fitness value, resulting in an allele's eventual fixation or loss due to drift.

While modern population genetics offers powerful tools for describing allele-frequency dynamics [11–13], many standard formulations are optimized for environments that are

effectively stationary or can be summarized by time-averaged selection coefficients. In such settings, phenotypic plasticity is often modeled as environmental variance rather than as a mechanistic driver of adaptation [14,15]. This can obscure the expressed phenotypic diversity and latent adaptive potential carried by a genotype [16–18]. Moreover, many microbes exhibit fuzzy genotype, phenotype mappings and facultative aggregation [19–22], pooling behavioral and epigenetic states and enabling nonvertical information transfer that lineage-centric analyses miss. These views have been robustly challenged by the Extended Evolutionary Synthesis (EES), which posits that developmental plasticity and nongenetic information transfer are primary engines of evolutionary change, capable of leading, rather than just following, genetic adaptation, treating nongenetic inheritance as a causal contributor to evolutionary change [23–28].

Despite the conceptual maturity of the EES [24,26], a disconnect remains between these theoretical advances and the computational models used to quantify them. These omissions matter empirically. High-level antibiotic resistance can appear within weeks; *Pseudomonas* tolerates colistin in < 10 days [29]; drug-resistant cancers emerge over a single therapy course [30,31]; mosquito populations acquire multi-insecticide resistance between spray cycles [32]. Under classical diffusion with realistic mutation rates and fixation probabilities, expected waiting times are orders of magnitude longer, unless one invokes implausible effective sizes or hyper-mutators. A more plausible route is cue-driven phenotypic exploration [33,34]: efflux up-regulation, membrane remodeling, persister formation, many conformations of proteins, and the chaperoning mechanisms involved in their enforcement [4]. These states arise plastically and may later

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

be hardwired by mutation via genetic assimilation [35,36]. Classic work has shown how lifetime learning can guide evolutionary search by effectively smoothing rugged fitness landscapes [37,38], but most models have not incorporated a falsifiable explanation of how this could be implemented mechanistically. Interpreted through active-inference principles, cells may use internal models to minimize prediction error about forthcoming states and deploy preparatory phenotypes when cues are reliable [15,39], as suggested by EES.

To turn these qualitative ideas into quantitative, falsifiable predictions, we introduce a dual framework combining computational modeling with experimental validation. We develop a formal model of Markovian, Bayesian Agents and test its central prediction in the single-celled microbial eukaryote *Capsaspora owczarzakii* using a programmable Smart Incubator conditioning assay.

Microbial eukaryotes are ideal for the tests outlined [40,41]: compact genomes, rich regulatory toolkits (signal transduction, chromatin remodeling, adhesion), rapid growth, precise editing [42], and compatibility with microfluidics and high-throughput phenotyping. *Capsaspora owczarzakii* (hereafter *Capsaspora*), a free-living close relative of animals (within the Holozoa clade), is especially suitable: encodes integrin/kinase signaling, calcium-dependent adhesion, extensive chromatin machinery [43–45], and exhibits facultative aggregation, enabling direct tests of collective information grouping and highlighting its regulatory complexity [19].

The notion that even unicellular organisms exhibit learning-like behavior has a long history, dating to Jennings’ careful studies of *Stentor* habituation and individual behavioral variation [46]. Modern basal cognition research has revived and extended this view, arguing that minimal cognitive capacities are widespread and evolutionarily ancient [47–50] and suggested mechanisms to utilize it [51]. Here we hypothesize that some single-celled eukaryotes can be usefully modeled as Markovian, Bayesian Agents (MBAs): organisms that infer latent environmental structure within a lifetime and use cues to guide phenotype choice. We formalize this hypothesis through an explicit null-model comparison with a nonlearning Blind Agent (BA) whose strategy evolves only across generations. *In silico*, we compare MBA and BA populations across controlled environmental regimes that vary in cue predictability, landscape structure, and plasticity cost, deriving falsifiable predictions such as an MBA-to-BA crossover as cue reliability declines. *In vitro*, we test a central prediction using a smart-incubator conditioning assay in *Capsaspora* and outline additional assays (e.g., habituation and threshold adaptation) enabled by open-source platforms described in the companion repository.

Together, this framework could be used to test when intra-lifetime inference outperforms mutation-only adaptation, delineates the boundary where plasticity collapses as cues degrade, and identifies conditions under which plastic exploration is genetically assimilated. By integrating lineage-resolved measurements with explicit null models, we aim to close the gap between rapid real-world adaptation and classical population-genetic expectations, and to clarify how cellular plasticity and facultative aggregation can scaffold the earliest steps toward multicellularity and form the basis of highly complex behavior.

II. RESULTS

This section presents the model, simulation results, and an experimental test of a key prediction. We first define the simulated environment and two agent architectures, a nonlearning Blind Agent (BA) and a learning Markovian, Bayesian Agent (MBA), and quantify when learning yields a fitness advantage. We find that the MBA advantage depends on cue predictability and reverses when environmental stochasticity exceeds a threshold ($\varepsilon \approx 0.2$). Finally, we report a conditioning assay in *Capsaspora* demonstrating improved survival under predictive cue, stress pairing.

A. The simulated environment

Agents inhabit a world defined by a deterministic, five-stage daily cycle through Hidden Environmental States (HES), $s \in \{0, 1, 2, 3, 4\}$. We term these states “hidden” because agents do not directly observe the true state; they receive only noisy temperature cues ($T_{\text{obs}} = T_{\text{true}} + \mathcal{N}(0, 0.2)$) and must infer the underlying environmental state, analogous to Hidden Markov Model inference. Each state is associated with a unique fitness landscape for three phenotypes representing distinct metabolic strategies: P_1 (nitrogen-exploitation specialist, optimal during resource abundance), P_2 (generalist with balanced resource allocation), and P_3 (preparatory state that enables anticipatory responses). Three phenotypes provide minimal complexity for nontrivial strategic diversity while remaining computationally tractable. The fitness landscape is designed to be nontrivial, requiring anticipatory action for maximal fitness. Specifically, a conditional fitness bonus is awarded in the “dusk” (HES 4) and “dawn” (HES 0) states only if a preparatory phenotype (P_3) was expressed in the preceding “afternoon” state (HES 3), as illustrated in Fig. 1.

B. Agent architectures

1. The blind agent (BA)

The BA is a null model whose behavior is dictated by its 452-bit genome. The genome decodes into a five-stage phenotype sequence (the *default* progression), a baseline transition probability p_{base} , and a temperature sensitivity factor C . Importantly, agents do not rigidly follow the encoded sequence: transitions between sequence positions occur stochastically based on P_{eff} [Eq. (1)], allowing agents to advance through their sequence at variable rates. An agent may remain at the same sequence position across multiple environmental stages (if P_{eff} is low) or advance rapidly through multiple positions within a single stage (if temperature changes trigger high P_{eff}). Selection thus acts not only on which phenotypes appear in the sequence but also on the timing dynamics of phenotype transitions. The effective transition probability is given by

$$P_{\text{eff}} = p_{\text{base}} \times [1 + C \times \text{norm}(\Delta T)], \quad (1)$$

where $\text{norm}(\Delta T)$ is the normalized change in temperature. The BA cannot learn; adaptation occurs only through mutation and selection across generations.

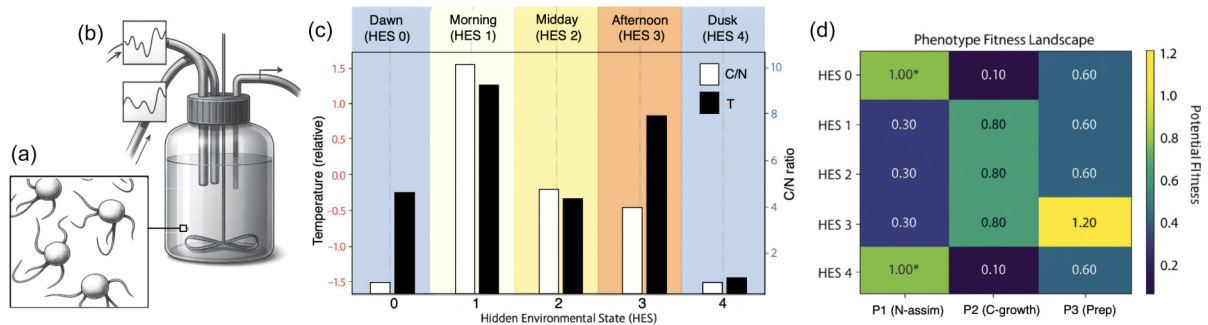


FIG. 1. The Simulated Environment and Fitness Landscape. (a) Biological system: *Capsaspora owczarzaki*, a unicellular holozoan serving as our model organism for experimental validation. (b) Cue-stress regimes (programmable in the Smart Incubator): schematic of the experimental control space, where cue modalities (e.g., light/vibration) can be paired to thermal stress under predictive vs nonpredictive temporal relationships, and schedule parameters (cycle durations, cue-stress lag) can be systematically varied. (c) Environmental dynamics: The simulation progresses through Hidden Environmental States (HES 0–4), with temperature (T , black bars) and Carbon-Nitrogen ratio ($C:N$, white bars) defining each state. (d) Fitness landscape: Each HES has unique fitness payoffs for phenotypes P_1 , P_2 , P_3 . The asterisk (*) indicates the conditional fitness bonus: maximum fitness in “dusk” (HES 4) and “dawn” (HES 0) requires prior expression of the preparatory phenotype (P_3) in “afternoon” (HES 3), using temperature change as a predictive cue. Panels (c) and (d) depict simulation parameters; experimental validation is presented below.

2. The Markovian Bayesian agent (MBA)

The MBA models a simple active inference agent with a dual-layer architecture:

- (1) Genome (prior): An innate, evolved strategy identical in structure to the BA's.
- (2) Learning layer (posterior): A mutable copy of the strategy that is updated through experience and drives behavior.
- (3) Cost of plasticity: A metabolic cost is incurred for divergence between the genomic prior and the learned posterior, calculated as

$$\text{Cost} = \lambda_H \cdot d_H(\mathbf{p}_g, \mathbf{p}_l) + \lambda_{KL} \cdot D_{KL}[P(p_{\text{base},g}) || P(p_{\text{base},l})], \quad (2)$$

where d_H is the Hamming distance between phenotype sequences and D_{KL} is the Kullback-Leibler divergence between transition probabilities. Both fitness and plasticity cost are dimensionless quantities: fitness is normalized to $[0,1]$ (reproductive success relative to maximum), and cost is scaled by coefficients $\lambda_H = 0.01$, $\lambda_{KL} = 0.02$ to remain small relative to fitness differences. The subtraction $F_{\text{net}} = F - C$ is thus well-defined and captures the trade-off between behavioral performance and metabolic investment in plasticity machinery.

(4) Regret-based learning: Poor performance triggers a “regret” signal that gates exploratory updating of the learned strategy. The optimal fitness $F_{\text{optimal}}(h)$ is used only to normalize this trigger (a calibrated stress threshold), not to provide gradient information about what to do next. When regret is high ($R > 0.2$) and a random draw succeeds ($U(0,1) < \eta$ with $\eta = 0.3$), the agent applies RANDSWITCH: choose a random sequence position $j \in \{1, \dots, 5\}$ and replace $L[j]$ with a uniformly random alternative from $\{P_1, P_2, P_3\} \setminus \{L[j]\}$. Thus, learning is undirected stochastic exploration gated by a stress-like signal; similar behavior would follow from using an absolute fitness threshold rather than an optimum-referenced regret.

(5) Genetic assimilation: A consistently successful learned strategy can be written back to the genome, becoming innate for offspring and reducing the agent's own plasticity cost.

C. Evolutionary dynamics

Populations evolve under a Moran process with a constant size of $N = 100$. An agent's total daily fitness is the average across the five states, with any plasticity costs for MBAs subtracted. Selection acts on this final fitness, with agents chosen to reproduce proportional to their success.

D. MBA outperforms BA in predictable environments

In separate simulations within a stable, deterministic environment, MBA populations consistently achieved higher fitness than BA populations [Fig. 2(a)]. Over 500 days, the mean fitness of the MBA populations stabilized at 0.692 ± 0.037 , a significant 31.8% advantage over the BA populations, which reached a fitness of only 0.525 ± 0.010 ($p < 3 \times 10^{-11}$). This advantage arises from the MBA's ability to learn the optimal phenotype sequence that anticipates environmental transitions. The population entropy of MBAs rapidly decreases as they converge on this optimal strategy, while BA entropy remains high, reflecting a diverse but less fit population [Fig. 2(b)]. The metabolic cost of plasticity for MBAs remained low, averaging just 0.014, indicating that the benefits of learning far outweighed the costs [Fig. 2(c)]. Müller plots of dominant genotypes confirm that MBA populations rapidly fix an optimal strategy, whereas BA populations maintain a diverse pool of suboptimal genotypes [Fig. 2(d)].

E. Environmental predictability defines the value of plasticity

To determine the limits of the MBA's advantage, we performed a stress test by introducing environmental stochasticity (ε). Here ε represents the probability that both the true environmental state and its perceived cue are indepen-

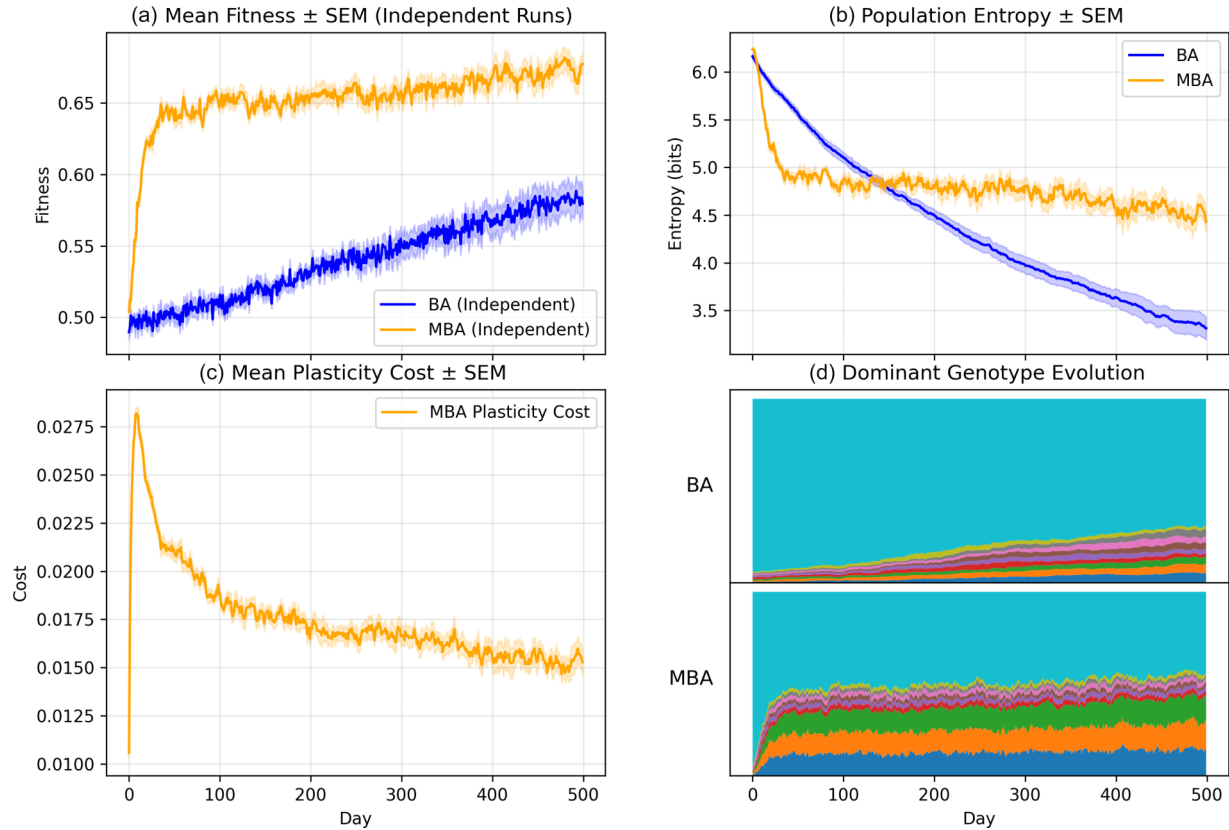


FIG. 2. Independent Population Dynamics. Comparison of BA-only and MBA-only populations over 500 days (mean $\pm$ SEM over 10 replicates). (a) MBA agents achieve significantly higher mean fitness. (b) MBA population entropy drops as agents converge on an optimal strategy, while BA entropy remains high. (c) The metabolic cost of plasticity for MBAs is minimal. (d) Müller plots show the rapid fixation of a dominant, optimal genotype in MBA populations (bottom), while BA populations (top) fail to converge.

dently randomized at each daily stage, breaking any learnable correlation. The results reveal a clear threshold where plasticity becomes disadvantageous [Fig. 3(a)]. MBAs maintain their fitness advantage for $\varepsilon \leq 0.15$. However, as stochasticity increases beyond $\varepsilon \approx 0.20$, the environment becomes too unpredictable to learn, and the metabolic cost of maintaining plasticity machinery renders the MBA less fit than the simpler, nonlearning BA. At maximum stochasticity ($\varepsilon = 1.0$), the BA's strategy of adopting a fixed, generalist phenotype is superior.

F. Plasticity confers resilience to abrupt environmental shifts

We next tested how populations respond to a sudden, permanent environmental change after a long period of stability (the “genetic lock-in” experiment). After 300 days of adaptation to a canonical environmental cycle, the sequence of daily states was permanently permuted. The MBA population experienced a severe, immediate fitness collapse, dropping by 39% on the first day postshift (from 0.699 to 0.426), illustrating the “cost of forgetting” where a previously optimal, learned strategy becomes maladaptive [Fig. 3(b)]. However, this deficit was transient. The MBAs’ plasticity allowed for rapid recovery, with fitness rebounding to 0.495 by day 2 and eventually stabilizing at 0.707, a level exceeding their preshift performance. In contrast, the BA population showed minimal response, with fitness remaining static around 0.515.

This experiment highlights the dual nature of plasticity: while it carries the risk of a temporary maladaptation when the environment changes, it provides the essential mechanism for rapid recovery and superior long-term adaptation.

G. MBAs’ adaptive advantage drives BAs to extinction in direct competition

To assess the ultimate evolutionary outcome, we simulated a large, mixed population of 500 BA and 500 MBA agents in direct competition for 100 000 days in a deterministic environment. The results are unequivocal (Fig. 4). Despite starting at equal numbers, the MBA population steadily outcompetes the BA population, driving it to complete extinction by approximately day 60 000. The MBA’s higher average fitness allows it to secure a greater share of reproductive opportunities in the Moran process. This long-term result confirms that the fitness advantages observed in isolated populations and short-term experiments translate into a decisive, long-term evolutionary victory for the plastic, learning-based strategy.

H. Quantifying drivers of adaptive advantage

Before dissecting parameter effects, we asked whether the MBA’s advantage might depend on the specific order of hidden environmental states (HES). In other words, is plasticity tied to the canonical daily cycle, or does it generalize across

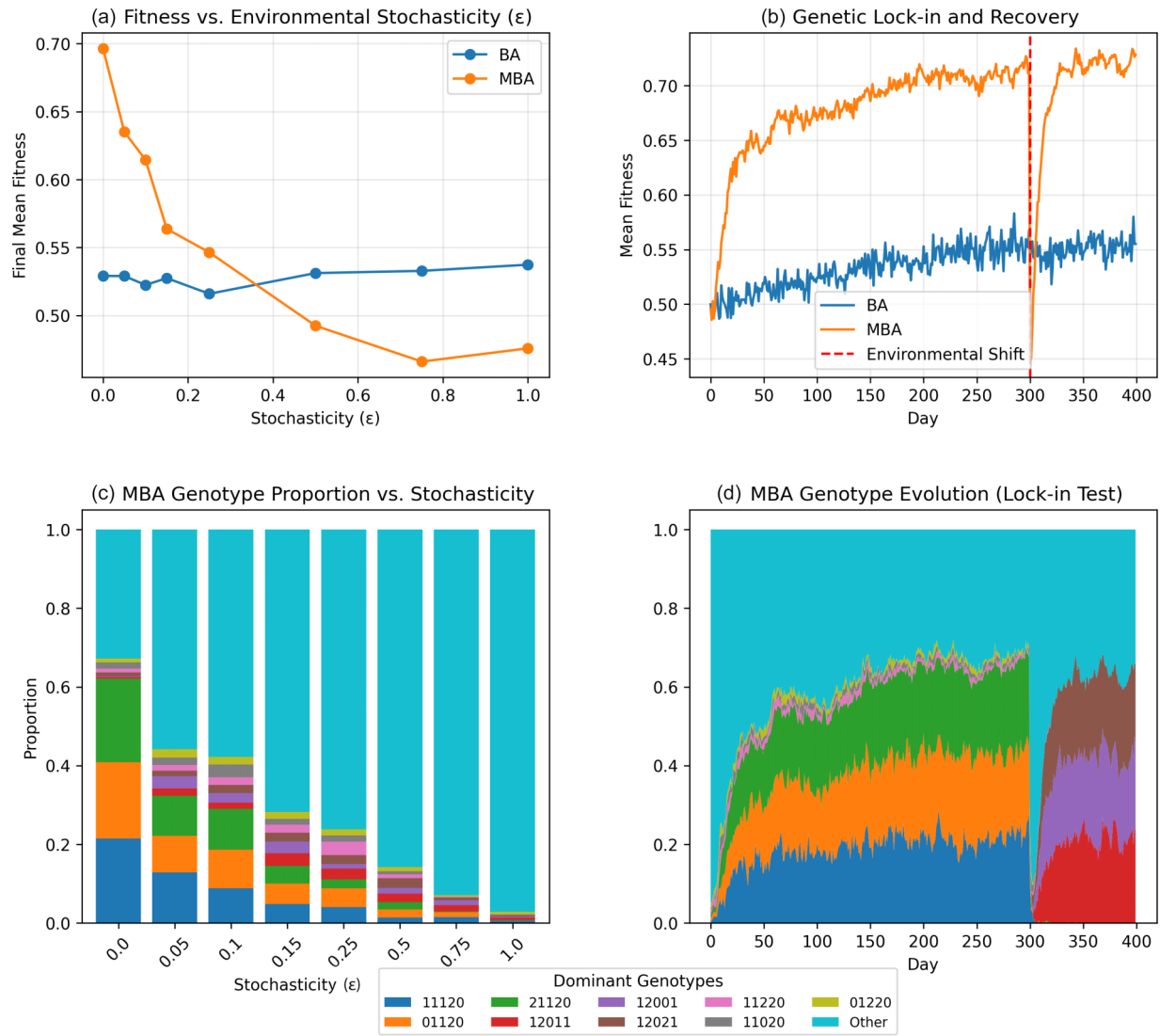


FIG. 3. Environmental Stress and Genetic Lock-in. (a) Fitness as a function of environmental stochasticity (ϵ). The MBA advantage erodes as the environment becomes unpredictable, with a crossover point at $\epsilon \approx 0.2$. (b) Response to a sudden environmental shift after 300 days. MBA fitness initially collapses (the “cost of forgetting”) but rapidly recovers to a level exceeding its preshift performance. BA fitness remains largely static. (c), (d) Müller plots for the MBA population during the stress test and lock-in experiment, respectively, showing the dynamics of genotype evolution.

arbitrary topologies? To this end, we exhaustively simulated 120 permutations of the five-state daily cycle (600 days, 100 agents), generating an ad hoc reward rule for each sequence. In all cases, the learning agent achieved a positive fitness advantage over the blind agent. No significant dependence on Hamming distance from the canonical sequence was detected (Kruskal-Wallis test, $H = 6.14$, $p = 0.19$), indicating that global sequence topology does not, by itself, abolish the advantage of learning. However, the magnitude of the advantage varied across permutations, suggesting that specific structural features of the sequence can modulate performance.

Having established that the learning strategy is robust to sequence order, we next quantified the relative importance of environmental and internal parameters using a large-scale parameter sweep (5000 independent runs, 600 days each). A CatBoost gradient boosting regression model was

trained to predict the learning agent’s fitness advantage (Δ , $\Delta_{\text{final_mean}}$). The model achieved high predictive accuracy (RMSE = 0.0277), allowing robust analysis of main effects and pairwise interactions.

Figure 5(a) shows the overall performance landscape across environmental stochasticity (ϵ) and internal regimes. A clear threshold is evident: the learning agent retains an advantage only at low stochasticity ($\epsilon < 0.2$); above this, fixed strategies outperform, as indicated by points falling below the red dashed line ($\Delta = 0$). Within the low-stochasticity regime, the highest fitness advantages are associated with low learning rates (small marker size) and low cost multipliers (darker colors).

To formally rank parameter contributions, we used SHAP (SHapley Additive exPlanations), a model-agnostic method that assigns additive importance values to features. As shown

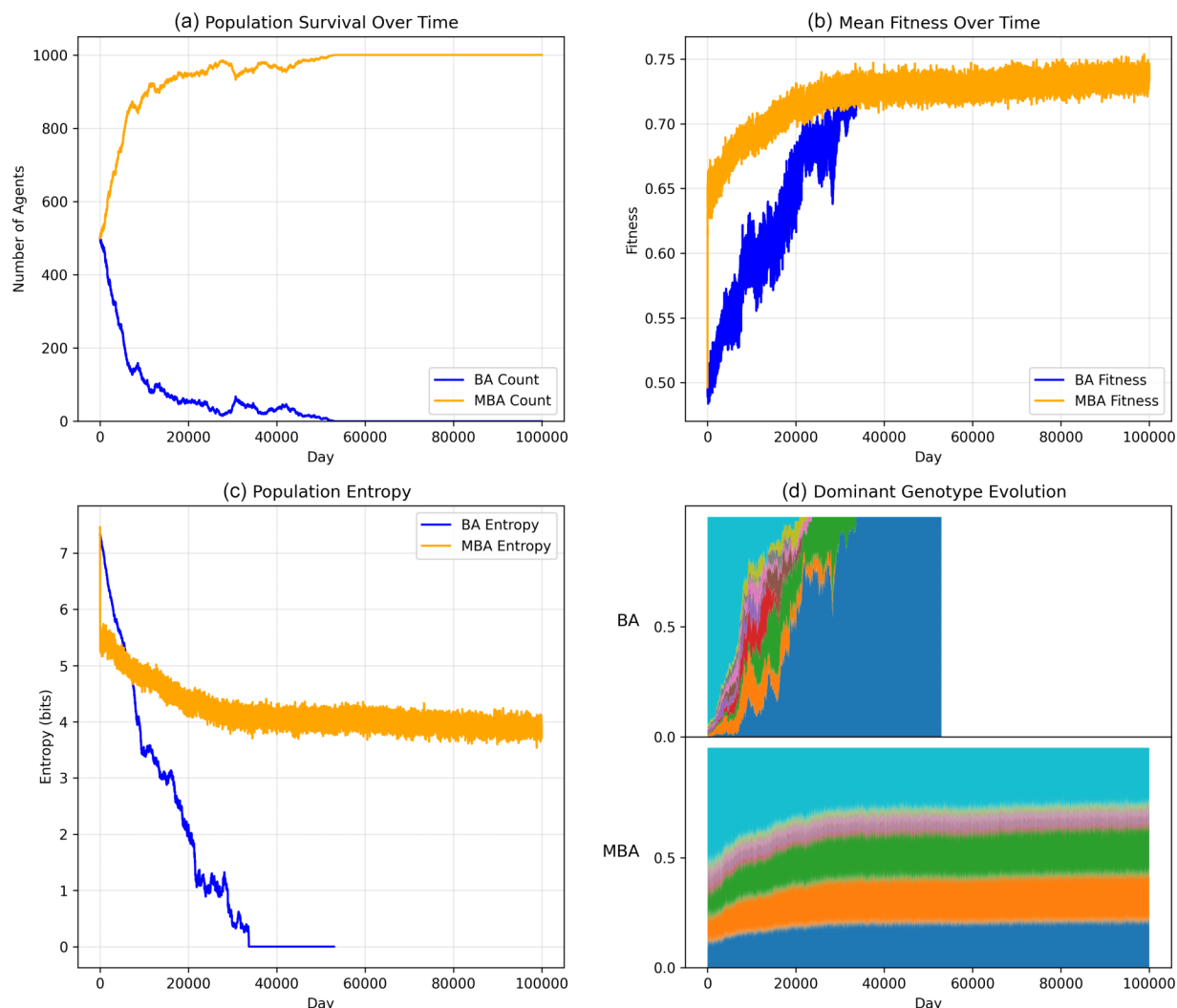


FIG. 4. Long-Term Direct Competition. In a mixed population starting with 500 BA and 500 MBA agents, the MBA population (orange) consistently outcompetes the BA population (blue), leading to the latter's extinction. (a) Population counts over time. (b) The mean fitness of the MBA population remains consistently higher than the BA population. (c) Population entropy for both agent types decreases as the MBA population expands and fixes its strategy. (d) Müller plots show the dominance and eventual fixation of the MBA genotype.

in Fig. 5(b), environmental stochasticity dominates all other predictors. Among the internal attributes, cost multiplier and learning rate are most influential, while environmental permutation (the structure of the hidden environmental sequence) also contributes substantially. Penalty exerts a smaller but consistent effect.

Finally, we quantified pairwise interactions using CatBoost's interaction strength metric. The resulting heatmap [Fig. 5(c)] shows that interactions are non-negligible. The strongest interaction is between stochasticity and learning rate, confirming that the effect of internal parameters depends on the environmental regime. The next strongest is between stochasticity and permutation, indicating that the tolerance of learning agents to noise depends on the structural predictability of the environment. Together, these analyses establish a hierarchy: environmental factors are the dominant drivers of adaptive advantage, internal parameters modulate outcomes, and their interactions determine the boundary conditions where learning is favored.

I. Experimental validation in *Capsaspora*

To test the model's central prediction, that the advantage of anticipatory responses depends on cue predictability, we performed a classical conditioning experiment in *Capsaspora* using a programmable Smart Incubator (see Materials and Methods). We compared a predictive condition ($\text{Corr} = 1$) to a nonpredictive control ($\text{Corr} = 0$) in which cue timing was randomized, then quantified survival after a lethal challenge.

1. Smart incubator: A classical conditioning paradigm for unicellular organisms

We designed this protocol utilizing the principles for forward (nonoverlap) Pavlovian conditioning and unpaired control groups outlined by Abramson and Levin [52]. The experimental design [Figs. 6(a)–6(c)] uses a neutral compound conditioned stimulus (CS; LED light and mechanical vibration) paired with a mild heat stress unconditioned stimulus (US; 32°C for 30 min) during a week-long training phase

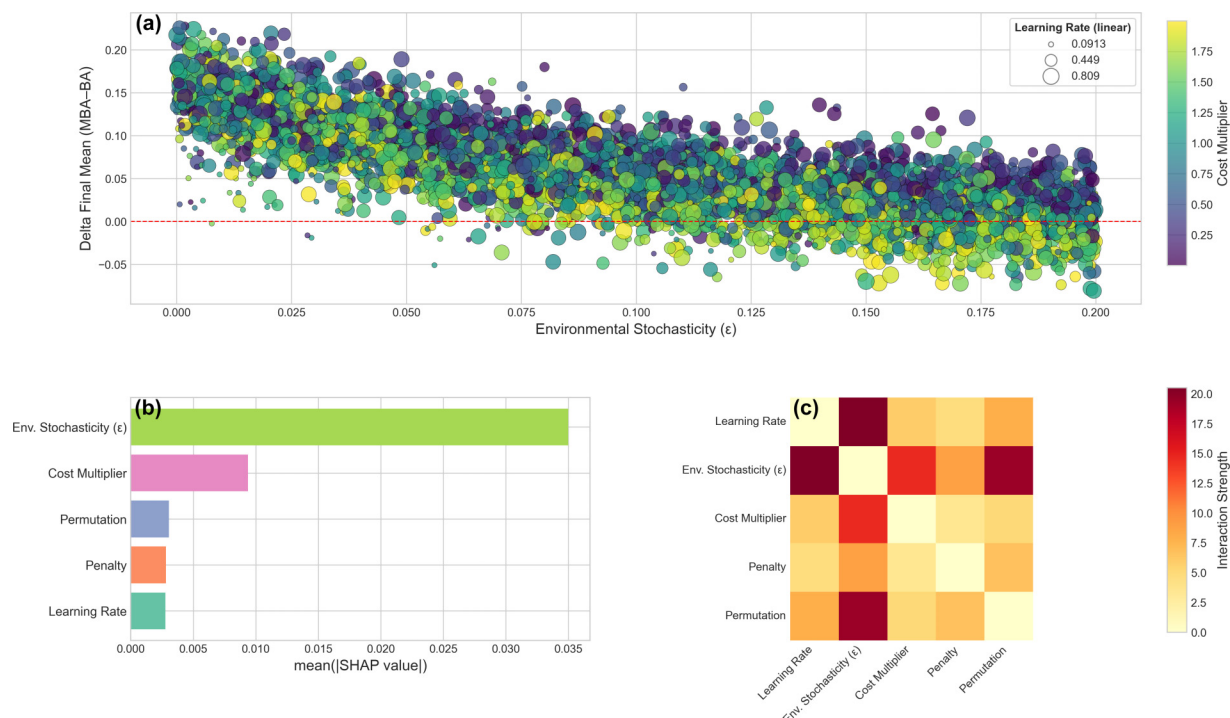


FIG. 5. Drivers of Adaptive Advantage. (a) Performance landscape of the learning agent across environmental stochasticity (ϵ), cost multiplier (color), and learning rate (marker size). The red dashed line marks $\Delta = 0$, the point where fixed strategies outperform. (b) Mean absolute SHAP (SHapley Additive exPlanations) values ranking the contribution of each parameter to the model's predictions. (c) Heatmap of pairwise interaction strengths, showing dependencies between environmental and internal parameters.

(~ 35 cycles). To reduce fatigue and sensory adaptation, we implemented long intertrial intervals and two distinct protocols:

(1) Predictive ($\text{Corr} = 1$): The CS reliably precedes heat stress with tight temporal coupling [CS–HS gap: 0.3–1.6 min; Fig. 7(d)]. This creates a learnable contingency [Fig. 6(a)].

(2) Nonpredictive ($\text{Corr} = 0$): The CS occurs at random times relative to heat stress [CS–HS gap: -245 to $+184$ min; Fig. 7(c)], destroying the predictive relationship while maintaining comparable cue exposure [Fig. 6(b)].

Following training, all cultures received a priming cue (full-strength CS at basal temperature) followed by a lethal heat challenge [38°C for 150 min; Fig. 6(c)]. Survival was assessed via propidium iodide (PI) staining and automated cell segmentation [Cellpose; Fig. 7(b)].

2. Experimental validation: Predictive cue-stress pairing enhances survival

The conditioning experiment yielded a striking result consistent with MBA predictions. Cultures trained in the predictive environment ($\text{Corr} = 1$) showed significantly lower mortality ($12.1 \pm 4.8\%$ dead cells) compared to cultures trained in the nonpredictive environment [$\text{Corr} = 0$; $26.8 \pm 5.2\%$ dead cells; Mann-Whitney U test, $p = 1.07 \times 10^{-8}$; Fig. 7(a)]. This represents an approximately 55% reduction in mortality for cells that experienced learnable cue-stress contingencies.

Representative microscopy [Fig. 7(b)] shows PI-stained cells with automated segmentation; dead cells (PI mean $> 2.5 \times 10^5$) are marked, and the PI intensity distribution

confirms clear separation of live and dead populations. The temporal dynamics of both protocols are shown in Figs. 7(c) and 7(d): individual temperature traces aligned to heat-shock onset reveal consistent cycle structure, while inset histograms confirm the critical difference—CS–HS gaps are tightly clustered near zero for the predictive condition but broadly distributed for the nonpredictive condition.

This result provides direct empirical support for the MBA framework: when environmental cues reliably predict upcoming stress, *Capsaspora* can exploit this contingency to mount anticipatory responses that enhance survival. When cue-stress relationships are randomized, no such advantage accrues, precisely as predicted by the ϵ threshold identified in our simulations.

III. DISCUSSION

Our results surface three recurring limits of diffusion-style, mutation-first accounts when environments are cue-rich and rapidly changing. First, a fixed genotype-phenotype map cannot capture latent internal states that let one lineage express multiple phenotypes without new mutations. Second, diffusion approximations calibrated to stationary or slowly varying regimes underrepresent high-frequency shocks unless explicitly modeled, while field surveys indicate that such shocks often dominate selection. Third, treating the environment as a purely external force ignores that the organism is an active participant. A genome does not just encode a static list of traits; it builds a dynamic control system [53] that must operate within strict physical boundaries. By neglecting this, classical models cannot account for the biophysical

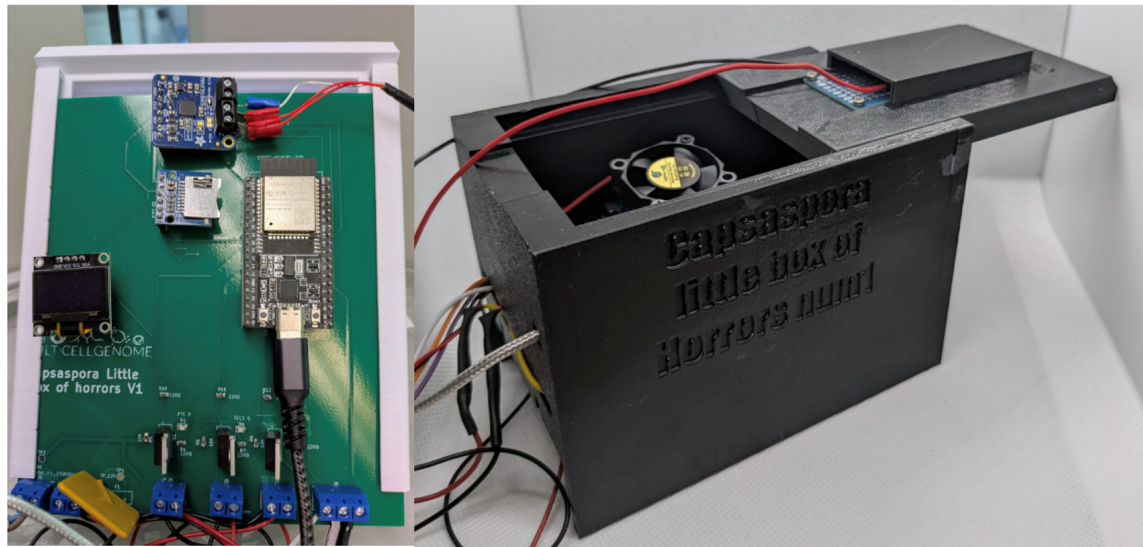
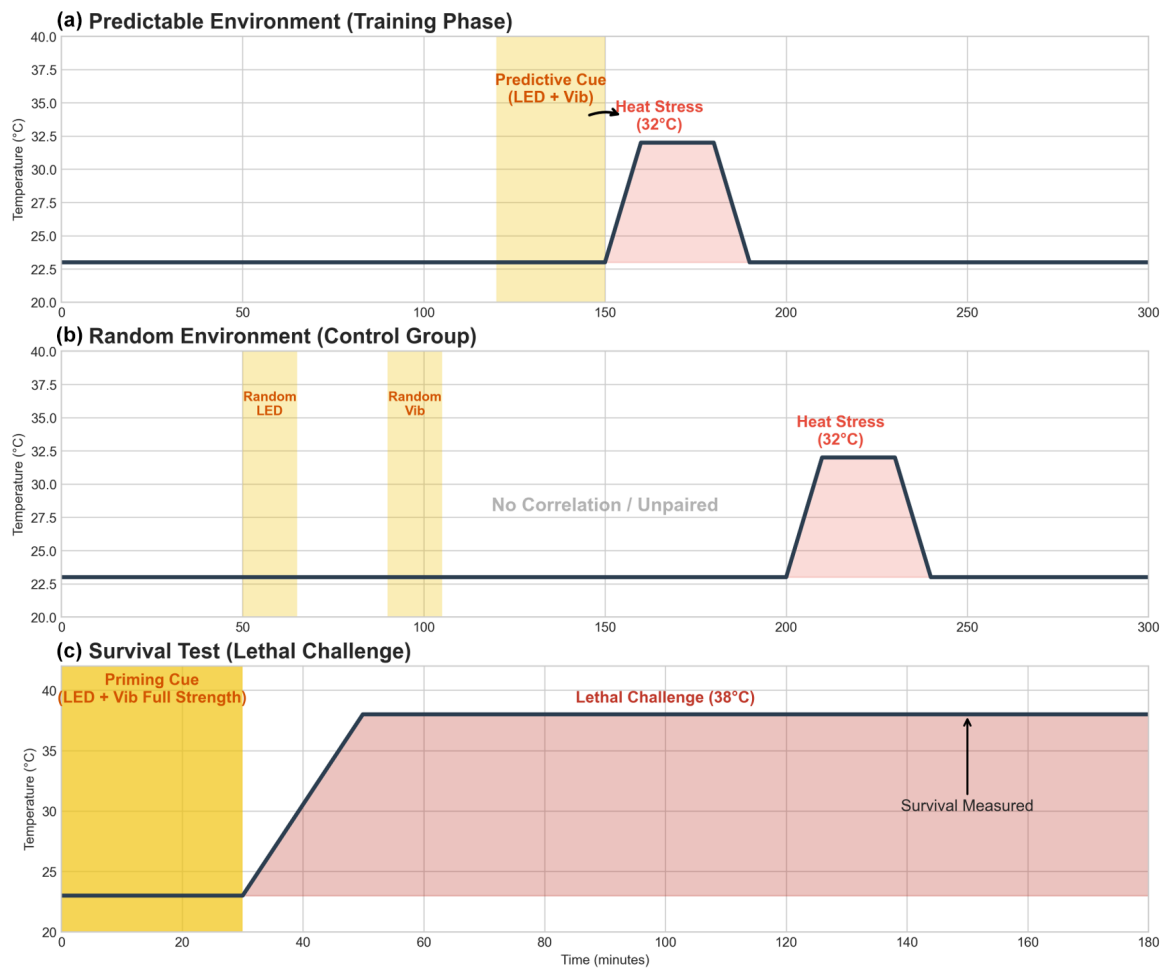


FIG. 6. Smart incubator: experimental paradigm and hardware. (a) Predictive environment (training phase): conditioned stimulus (CS; LED + vibration) reliably precedes heat stress (32 °C), creating a learnable cue-stress contingency. (b) Random environment (control): CS timing is randomized relative to heat stress, destroying predictability. (c) Survival test: after training, all cultures receive a priming cue followed by a lethal heat challenge (38 °C, 150 min); survival is measured via PI staining. (d) PCB control board featuring ESP32 microcontroller, thermal sensors, and solid-state relays for automated cue-stress delivery. (e) Custom 3D-printed incubator enclosure with Peltier heating/cooling elements.

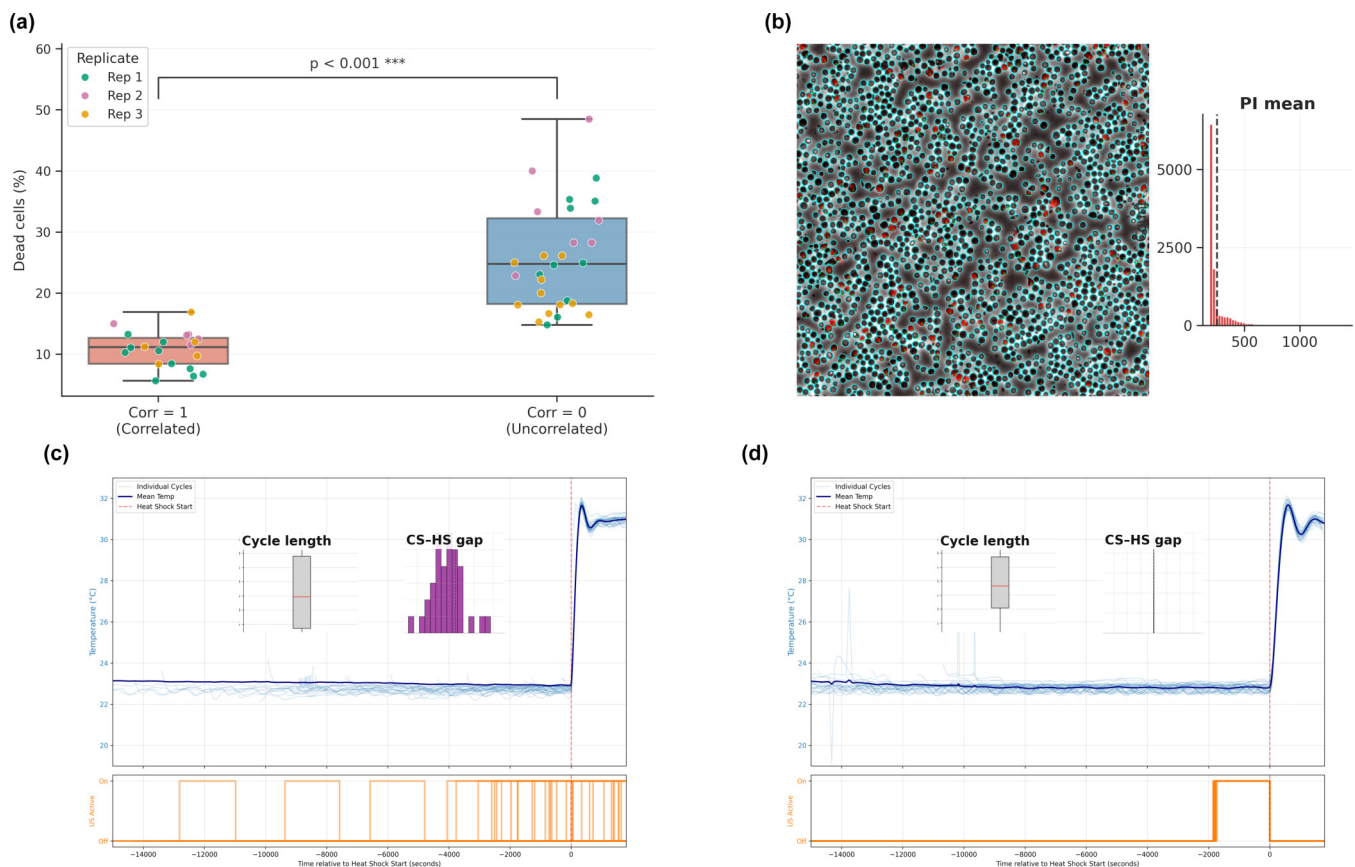


FIG. 7. Experimental validation of cue-contingency learning in *Capsaspora*. (a) Dead-cell fraction per image for correlated (Corr = 1) vs uncorrelated (Corr = 0) stress protocols, shown as boxplots with individual image points. Cultures trained with predictive cue-stress pairing show significantly reduced mortality (Mann-Whitney U , $p = 1.07 \times 10^{-8}$). (b) Representative microscopy field with Cellpose segmentation (blue outlines); dead cells marked with green X, live cells with red V; PI mean threshold = 2.5×10^5 . Right panel shows the PI intensity distribution. (c) Nonpredictive protocol (Corr = 0): temperature traces aligned to heat-shock (HS) onset, with CS activity shown below. Inset histograms show cycle-length distribution and CS-HS gap; gaps are broadly distributed (−245 to +184 min). (d) Predictive protocol (Corr = 1): same layout as (C); CS-HS gaps are tightly clustered near zero (0.3–1.6 min), confirming temporal coupling between cue and stress.

constraints (e.g., the speed of transcription, the energy cost of plasticity) and information-processing limits that ultimately determine how and when an organism can adapt. These omissions may be inconsequential in simplified laboratory settings, where learnable cues are scarce, and organisms behave effectively as Blind Agents (BAs), but we claim they become critical in natural, cue-rich ecologies.

Our agent-based simulations clarify when Markovian–Bayesian Agents (MBAs), plastic, model-updating organisms, gain or lose an advantage. In highly stochastic regimes (environmental noise ε above a critical threshold; Fig. 3), the predictive benefit collapses: metabolic and error costs dominate, and lower-cost BAs are more robust. In contrast, large deterministic shifts (e.g., a permanent change in cycle structure; Fig. 3) ultimately select for MBAs. Although MBAs initially lose fitness, they relearn and recover within tens of generations, whereas BAs require hundreds of generations for rare favorable mutations to arise and fix. This aligns with ecological patterns: short, unpredictable disturbances (storms, pulses) favor BA-like strategies (dormancy, spores), while sustained, directional pressures (chronic antibiotics, gradual climate shifts) favor plastic strains that can anticipate and precondition.

A notable limitation of our MBA architecture is that agents cannot learn to *stop* learning; the plasticity machinery and its associated costs are genomically fixed. In highly stochastic environments where no learnable pattern exists ($\varepsilon \rightarrow 1$), an ideal agent might benefit from recognizing this and adopting a fixed bet-hedging strategy. Our MBA cannot do this within its lifetime; instead, the population must evolve reduced plasticity costs across generations. This is biologically realistic: the sensory systems, signal transduction pathways, and regulatory networks that enable plasticity represent fixed developmental investments. Organisms in chronically unpredictable environments are observed to evolve reduced plasticity machinery rather than learn to ignore their sensors [54]. This constraint generates a testable prediction: the $\varepsilon \approx 0.2$ threshold we identify should correspond to an evolutionary boundary above which selection favors genomic reduction of plasticity investment.

An important extension of this framework is to compare MBAs against alternative null models that implement memory through dynamical internal states rather than explicit strategy edits. Such dynamical-systems agents would relax the fixed genotype-phenotype mapping while remaining devoid of a structural plasticity mechanism, providing a sharper test of

whether our inferred advantages depend on true learning or merely on dynamical continuity. We view this as a natural next step.

Our experimental results with *Capsaspora* provide direct empirical support for these simulation predictions. In the conditioning experiment, cultures exposed to predictive cue-stress pairing ($\text{Corr} = 1$) showed 55% reduced mortality compared to cultures exposed to nonpredictive, randomized cue timing ($\text{Corr} = 0$; $p = 1.07 \times 10^{-8}$; Fig. 7). This maps directly onto the ε threshold identified in simulations: when the cue-stress correlation is high ($\text{Corr} = 1$, analogous to low ε), MBA-like anticipatory strategies confer a substantial fitness advantage; when the correlation is destroyed ($\text{Corr} = 0$, analogous to high ε), no such advantage accrues. The magnitude of the experimental effect, approximately 55% mortality reduction, is consistent with the 30%–40% fitness advantages observed in simulations under comparable conditions, providing quantitative as well as qualitative validation of the framework.

Beyond short-term adaptation, long-run competition reveals a buffering role for plasticity we term the *chaperone effect* (Fig. 4). MBA populations maintain an acquired “posterior” layer that cushions the genomic “prior” from deleterious mutations: individuals can express high-fitness phenotypes despite imperfect genotypes, preserving adaptive memory at the population level. Operationally, this appears as sustained anticipatory heterogeneity; plasticity costs prevent a single genotype from sweeping, stabilizing population entropy at high values, whereas BA populations converge to a narrow genotypic/phenotypic state before collapsing by drift. Conceptually, this connects to classic ideas of plasticity-first evolution and genetic capacitors (e.g., Baldwin, Waddington, Hsp90), but here emerges from explicit online inference with quantified entropy trajectories. Intriguingly, this buffering role of plasticity mirrors what is now emerging in cancer evolution, where nonheritable phenotypic plasticity precedes and scaffolds genetic stabilization. Recent single-cell studies argue that cancer progression, metastasis, and drug resistance often follow a phenotypes-first blueprint, in which stress-induced plastic states provide the raw material for later genetic fixation [31,55].

This logic also suggests an evolutionary route toward simple multicellularity during Proterozoic instability: physical compartmentalization can shelter interior tissues from an otherwise high-entropy exterior, converting an unpredictable environment into a locally predictable micro-niche in which costly, MBA-like inference becomes evolutionarily stable. Importantly, plasticity may not be as metabolically extravagant as traditionally assumed. Work on cognitive energetics shows that active cognition in neural systems increases energy use by only $\sim 5\%$ over baseline homeostasis, yet yields disproportionately large adaptive returns. If cellular cognition scales similarly, then the return-on-information (ROI) may be immense: modest energetic investments in inference can yield orders-of-magnitude gains in survival [56] and adaptability [50,57].

Our results further highlight that environmental perturbations are the dominant triggers for learning. Crucially, their impact is not absolute but species-relative: filtered through doubling time, sensory repertoire, and life cycle. What manifests as high-frequency noise for one organism

may constitute stable structure for another. Our experimental platforms are designed precisely to capture this relativity. The Smart Incubator and Microscope-Mounted Chemostat generate structured environments with tunable rates of change, cue reliabilities, and stress couplings, creating tailored “worlds” to test how different organisms perceive and act within them. In this sense, the platforms serve not merely as tools but as controlled ecologies for recapitulating the axes of our framework, hidden environmental states (HES), environmental noise (ε), and temporal coupling, in a testable, reproducible way.

Finally, while many aspects of our framework are anchored in established biology, learning rates corresponding to stress-induced phenotypic changes, costs to metabolic and phenotypic overhead, and updating priors using genetic assimilation and the Baldwin effect, our work also isolates a deeper organizing principle: the equivalent of a *loss function* at the cellular level. We argue that in order to sustain an internal, updateable world model, a cell must encode not only prediction error but also the subjective weight of “how bad” different mismatches feel. This capacity for cellular “regret” is the missing substrate for closed-loop inference, bridging open-loop reaction with adaptive world modeling. While this may seem abstract, plausible molecular substrates for such durable state weighting are emerging. For instance, self-perpetuating prion-like protein assemblies can create long-term cellular memories of transient stresses, effectively acting as a physical instantiation of a loss function by encoding the persistent “cost” of a past prediction error [58]. Integrating these components together, our framework not only connects plasticity, entropy, and assimilation into a unified conceptual model but also identifies cellular regret as the critical principle to be empirically pursued.

This active inference solution-motivated search warrants a genome-scale interpretation: regret-gated exploration should be most effective when the actionable search space is small, and the effective landscape is relatively smooth. As genomes and fitness landscapes grow (providing many local optima as stopping points), selection can instead favor changes in higher-level regulation that co-modulate large sets of genes, effectively coarse graining (binning) the exploratory space and enabling rapid stress alleviation with fewer exploratory degrees of freedom. Repeated bouts of such coarse-grained adaptation can then be refined and genetically assimilated into more specific, gene-level solutions [59]. In this view, the model’s crossover behavior is not tied to “genome length” in a literal sense, but to the dimensionality of the control policy available to an organism at the timescale of plastic updating.

Beyond evolutionary inference, this perspective opens a new horizon for synthetic biology. Current engineering approaches often rely on supervised-learning analogs, designing specific genetic circuits to execute known tasks under fixed rules. If organisms can instead operate as MBAs, capable of unsupervised learning by tying internal regulatory adjustments to real-time fitness outcomes, then a new design paradigm becomes possible. Instead of scripting solutions, one could scaffold internal loss functions and let cells discover functional strategies through inference and selection. This would enable the engineering of cells that not only execute logic, but actively explore, anticipate, and generalize, an

TABLE I. Environmental state parameters.

HES	Stage	Temperature (T)	Carbon (C)	Nitrogen (N)	C:N Ratio
0	Dawn	-0.34	0.1	1.0	0.1
1	Morning	1.38	1.0	0.1	10.0
2	Midday	-0.34	0.8	0.2	4.0
3	Afternoon	0.80	0.8	0.2	4.0
4	Dusk	-1.49	0.1	1.0	0.1

adaptive architecture closer to evolutionarily refined cognition than mechanistic control.

IV. MATERIALS AND METHODS

A. Simulation environment

The simulation models a population of agents evolving in a cyclical environment with five discrete Hidden Environmental States (HES 0, 4). Each day consists of a deterministic progression through all five states in sequence: HES 0 → HES 1 → HES 2 → HES 3 → HES 4, then cycling back to HES 0 on the following day. Each environmental state is characterized by three parameters: normalized temperature (T), carbon concentration (C), and nitrogen concentration (N), as shown in Table I.

At each environmental state transition, agents receive noisy observations of the true environmental parameters, with Gaussian noise ($\sigma = 0.2$) added to each dimension. To test robustness, we implemented an environmental stochasticity parameter (ε) that introduces a probability of random state transitions instead of the deterministic sequence.

B. Agent architecture

Both agent types share an identical 452-bit boolean genome structure that encodes three key components:

1. Genome structure

(1) Phenotype sequence (320 bits): Five 64-bit vectors, each encoding one of three possible phenotypes (P1, P2, or P3) using a nearest-centroid decoding scheme with predefined centroids:

(1) P1 centroid: All zeros (64 bits of 0)

(2) P2 centroid: Half ones, half zeros (32 bits of 1, 32 bits of 0)

(3) P3 centroid: All ones (64 bits of 1)

(4) Decoding: Each 64-bit vector is assigned to the closest centroid via Hamming distance

(2) Baseline transition probability (100 bits): Encodes a probability $p_{\text{base}} \in [0, 1]$ calculated as the proportion of "1" bits

(3) Temperature sensitivity (32 bits): Encodes a sensitivity factor $C \in [-1, 1]$ normalized around the midpoint: $C = \frac{\text{num_ones} - 16}{16}$

2. Phenotype transition mechanics

Agent phenotype transitions are governed by a temperature-sensitive probabilistic mechanism. The effective

TABLE II. Base fitness matrix.

HES	P1	P2	P3
0	1.0	0.1	0.6
1	0.3	0.8	0.6
2	0.3	0.8	0.6
3	0.3	0.8	1.2
4	1.0	0.1	0.6

transition probability at each environmental state is calculated as

$$P_{\text{effective}} = p_{\text{base}} \times [1 + C \times \text{norm}(\Delta T_{\text{obs}})], \quad (3)$$

where ΔT_{obs} is the observed temperature change between consecutive states, and $\text{norm}()$ normalizes the temperature change to $[0, 1]$. This mechanism allows agents to use temperature cues to time their phenotype transitions appropriately within the daily cycle.

C. Fitness landscape

The fitness landscape implements a complex multipeak structure with conditional bonuses designed to create valley-crossing challenges. The base fitness matrix assigns different fitness values to each phenotype in each environmental state, as shown in Table II.

1. Conditional fitness mechanism

A critical feature of the fitness landscape is the preparatory phenotype mechanism. When an agent expresses phenotype P3 in HES 3, it activates a two-step "preparation countdown" that enables maximum fitness (1.0) for phenotype P1 in subsequent nitrogen-rich states (HES 4 and HES 0). Without this preparation, P1 expression in nitrogen-rich states yields only 30% of its potential fitness (0.3), creating a temporal dependency that requires predictive adaptation.

D. Mathematical model of fitness dynamics

1. Formal model for fitness gain

We can formalize the fitness dynamics for both agent types as follows.

Blind agent (BA) fitness model. For a BA agent i at time t , the fitness is determined solely by its genetically encoded phenotype sequence:

$$V_{\text{BA},i}(t) = \frac{1}{5} \sum_{h=0}^4 F(G_i(t)[\phi_i(h)], h) - P_{\text{conditional}}(G_i(t), h), \quad (4)$$

where

(1) $G_i(t)$ is the genetically determined phenotype sequence at generation t

(2) $\phi_i(h)$ is the phenotype index used in environmental state h

(3) $F(p, h)$ is the base fitness lookup function for phenotype p in state h

(4) $P_{\text{conditional}}$ accounts for the preparatory phenotype penalty (0.7 reduction if P1 used without P3 preparation)

The evolutionary dynamics follow

$$G_i(t+1) = G_i(t) \oplus M(\mu), \quad (5)$$

where $M(\mu)$ represents bit-flip mutations with probability $\mu = 10^{-4}$.

Markovian–Bayesian agent (MBA) fitness model. For an MBA agent i at time t , fitness incorporates both learned strategy and plasticity cost:

$$V_{\text{MBA},i}(t) = \frac{1}{5} \sum_{h=0}^4 (F(L_i(t)[\phi_i(h)], h) - C_{\text{plasticity}}(L_i(t), G_i(t)) - P_{\text{conditional}}(L_i(t), h)), \quad (6)$$

where

- (1) $L_i(t)$ is the learned phenotype sequence (posterior)
 - (2) $G_i(t)$ is the genomic phenotype sequence (prior)
 - (3) $C_{\text{plasticity}}$ is the continuous metabolic cost of plasticity
- The plasticity cost function is defined as

$$C_{\text{plasticity}}(L_i, G_i) = \lambda_H \sum_{j=1}^5 I(L_i[j] \neq G_i[j]) + \lambda_{KL} D_{KL}(p_{\text{learned}} \| p_{\text{genomic}}), \quad (7)$$

where $I()$ is the indicator function, $\lambda_H = 0.01$, and $\lambda_{KL} = 0.02$.

Learning dynamics. The MBA's learned strategy evolves within lifetime according to

$$L_i(t, h+1) = \begin{cases} L_i(t, h) \oplus \text{RANDSWITCH}, & \text{if } R(t, h) > 0.2, \\ & \text{and } U(0, 1) < \eta, \\ L_i(t, h), & \text{otherwise,} \end{cases} \quad (8)$$

where $R(t, h) = F_{\text{optimal}}(h) - F_{\text{realized}}(t, h)$ is the regret function and $\eta = 0.3$ is the learning rate.

When an MBA achieves sustained success, learned adaptations can be genetically assimilated:

If $\sum_{\tau=t-K+1}^t I(V_{\text{MBA},i}(\tau) > \theta) \geq K$, then $G_i(t+1)[j] \leftarrow L_i(t)[j]$, where $K = 3$ consecutive successful generations and $\theta = 0.65$ is the fitness threshold.

Comparative advantage model. The relative advantage of MBA over BA can be expressed as

$$\Delta(t) = E[V_{\text{MBA}}(t)] - E[V_{\text{BA}}(t)]. \quad (9)$$

This advantage is positive when the learning benefit outweighs the plasticity cost:

$$E[F(L(t)) - F(G(t))] > E[C_{\text{plasticity}}(L(t), G(t))]. \quad (10)$$

The model predicts MBA advantage in predictable environments where learning can identify optimal strategies faster than mutation alone, but BA advantage in highly stochastic environments where plasticity costs exceed learning benefits.

E. Agent types

1. Blind agent (BA)

The Blind Agent represents a purely genetic strategy with no learning capability. Its behavior is entirely determined by its decoded genome:

- (1) Phenotype sequence: Fixed at birth from genetic encoding
- (2) Transition timing: Based solely on genetically encoded temperature sensitivity and baseline probability
- (3) No plasticity cost or learning mechanism
- (4) Relies entirely on mutation ($\mu = 10^{-4}$ per bit) and selection for adaptation

2. Markovian–Bayesian agent (MBA)

The MBA implements a dual-layer architecture enabling within-lifetime learning:

Dual-layer strategy.

- (1) Genomic layer (prior): The inherited 452-bit genome representing the evolutionary baseline
- (2) Learning layer (posterior): A mutable copy of the phenotype sequence and transition probability that drives actual behavior

Plasticity cost. The MBA pays a continuous metabolic cost for any divergence between its learned strategy and genomic prior:

$$C_{\text{plasticity}}(L_i, G_i) \quad [\text{see Eq. (7)}], \quad (11)$$

where

- (1) Hamming component: $\lambda_H = 0.01$ per differing phenotype position
- (2) KL divergence component: $\lambda_{KL} = 0.02$ for Bernoulli transition probabilities; clamp with $\varepsilon = 10^{-6}$ for numerical stability
- (3) Bernoulli KL formula:

$$D_{KL}(p \| q) = p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q}.$$

Learning mechanism. The MBA employs a regret-based learning system operating on daily performance windows:

- (1) Daily regret calculation:
 - (a) Optimal fitness computed as maximum possible fitness across all phenotypes and preparation states for each HES
 - (b) Daily regret = mean optimal fitness – mean realized fitness across all HES states in the day
- (2) Learning trigger: Learning occurs when both conditions are met:
 - (a) Daily regret > 0.2 (regret threshold)
 - (b) Random draw $< \eta$ (learning rate, default $\eta = 0.3$)
- (3) Phenotype update: When triggered, learning randomly:
 - (a) Selects one position in the five-stage phenotype sequence
 - (b) Replaces current phenotype at that position with random alternative (P1, P2, or P3)
- (4) Transition probability update: Gradual adjustment toward optimal target:
 - (a) $p_{\text{learned}} \leftarrow p_{\text{learned}} + \eta_p \times (0.9 - p_{\text{learned}})$
 - (b) Where $\eta_p = 0.05$ is the probability learning rate

Genetic assimilation. Whole-genome assimilation implements the Baldwin effect through sustained success tracking:

- (1) Trigger condition: Agent maintains high fitness (> 0.65) for $K = 3$ consecutive days
- (2) Assimilation process: Entire learned strategy (phenotype sequence + transition probability) written back to genome
- (3) Bit-level encoding: Learned phenotypes encoded as corresponding centroids, transition probability as shuffled bit vector
- (4) Cost reduction: Postassimilation, learned layer equals genomic layer, eliminating future plasticity costs
- (5) Heritability: Assimilated traits transmitted to offspring through normal reproduction

F. Population dynamics

Evolution proceeds via a Moran process maintaining constant population size:

- (1) Parent selection: Agents are selected for reproduction with probability proportional to their daily fitness
- (2) Victim selection: A random agent is chosen for replacement
- (3) Offspring production: The parent produces a clonal offspring subject to mutation
- (4) Mutation: Each genome bit flips with probability $\mu = 10^{-4}$ (identical for both BA and MBA agent types)

G. Experimental protocols

All experiments are orchestrated through a standardized wrapper system that ensures reproducible execution, consistent output formats, and systematic data collection. The wrapper architecture separates experiment orchestration from core simulation logic, enabling systematic parameter sweeps and comparative analysis.

1. Wrapper architecture

The experimental framework consists of five main wrapper types, each targeting specific research questions.

1. Vanilla wrapper (`vanilla.py`). Purpose: Baseline MBA vs BA comparison validation

- (1) Configuration: 1000 days, 10 replicates, 100 agents per type
- (2) Execution: Independent MBA-only ($n_{mba} = 100$, $n_{ba} = 0$) and BA-only ($n_{mba} = 0$, $n_{ba} = 100$) populations
- (3) Parameters: Learning rate $\eta = 0.3$, cost multiplier $= 1.0$, $\varepsilon = 0.0$, penalty $\gamma = 0.7$, canonical permutation (0,1,2,3,4)
- (4) Output: Delta analysis CSV, time series plots, population entropy trajectories
- (5) Validation: Sanity gates require $\Delta_{final_mean} \geq 0.10$ and $proportion_positive \geq 0.7$

2. Stress test wrapper (`stress.py`). Purpose: Environmental stochasticity threshold identification

- (1) Configuration: 1000 days, 10 replicates per epsilon level
- (2) Epsilon sweep: [0.01, 0.1, 0.2, 0.3, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0]
- (3) Execution: MBA vs BA comparison at each epsilon level

- (4) Expected outcome: MBA advantage decreases monotonically with increasing ε , crossing zero at $\varepsilon \approx 0.2$

3. Topology scan wrapper (`topology.py`). Purpose: Robustness across environmental permutations

- (1) Configuration: 600 days, five replicates, all 120 permutations of HES sequence
- (2) Metrics: Hamming distance to canonical cycle, P3 slot position, P1 slot distribution
- (3) Analysis: MBA advantage vs topological features (Kruskal, Wallis test across Hamming distances)

4. Grid sweep wrapper (`grid_sweep.py`). Purpose: Multidimensional parameter space exploration

- (1) Configuration: Factorial design across four parameters $\times$ five permutations
 - (a) Epsilon: [0.0, 0.1, 0.2, 0.25]
 - (b) Learning rate: [0.1, 0.3, 0.5, 0.7]
 - (c) Cost multiplier: [0.5, 1.0, 1.5, 2.0]
 - (d) Penalty: [0.2, 0.5, 0.7, 0.8]

- (2) Scale: 320 parameter combinations, 600 days, five replicates each

(3) Parallelization: Configurable worker threads for concurrent execution

(4) Output: `grid_summary.csv` with per-cell delta statistics, heatmap visualizations

5. Continuous sweep wrapper (`continuous_sweep.py`). Purpose: High-density sampling for regression analysis

- (1) Configuration: Randomized parameter sampling within ranges
 - (a) Epsilon: [0.0, 0.2], Learning rate: [0.0, 0.9]
 - (b) Cost multiplier: [0.0, 2.0], Penalty: [0.0, 0.8]
- (2) Scale: 5000+ experiments, random permutation selection
- (3) Format: `continuous_summary.csv` compatible with CatBoost/SHAP analysis
- (4) Incremental: Supports resumable data collection for ongoing analysis

2. Unified driver system

All experiments execute through a single unified driver (`unified_driver.py`) that ensures consistent:

- (1) Genome initialization: 452-bit boolean vectors with deterministic seeding
- (2) Population dynamics: Moran process with fitness-proportional selection
- (3) Mutation rate: $\mu = 10^{-4}$ per bit across both agent types
- (4) Output format: Standardized CSV schema with agent-level and population-level metrics
- (5) Reproducibility: Explicit seed management for deterministic replication

3. Data collection standards

Agent-level records. Each agent record captures:

- (1) Identity: `rep_id`, `day`, `agent_id`, `agent_type`
 - (2) Phenotype: `phenotype_sequence` (five-stage), `daily_fitness`, `plasticity cost`
 - (3) Genome: `genome_hash`, `genotype string`, `age`
 - (4) Learning: `learning_events`, `preparation_success_rate`
- Population-level metrics.*

(1) Entropy: Shannon entropy of phenotype sequence distribution

(2) Fitness statistics: Mean, variance, percentiles across agent types

(3) Genetic diversity: Number of unique genotypes, fixation events

Delta analysis. Standardized MBA-BA comparison:

(1) Window-based metrics: Final 200-day average for convergence assessment

(2) Statistical testing: Mann-Whitney U tests, proportion of positive deltas

(3) Visualization: Time series, distribution plots, effect size estimation

4. Computational implementation

(1) Language: Python 3.9+ with NumPy vectorization

(2) Architecture: Modular package structure (mba_vs_ba_sim/)

(3) Parallelization: joblib-based parallel execution for parameter sweeps

(4) Scalability: Memory-efficient streaming for large-scale experiments

(5) Validation: Integrated sanity gates and consistency checks

(6) Documentation: Comprehensive manifest files tracking all experimental parameters and outcomes

5. Analysis pipeline

Machine learning integration. Large-scale parameter sweeps integrate directly with ML analysis:

(1) CatBoost regression: Trained on continuous_summary.csv to predict $\Delta_{\text{final_mean}}$ from parameters

(2) SHAP analysis: Model-agnostic feature importance and interaction effects

(3) Performance: Typical RMSE < 0.03 on 5000+ experiments with cross-validation

Visualization framework. Standardized plotting utilities generate publication-ready figures:

(1) 3D Parameter Landscapes: Interactive scatter plots with size $\propto \Delta$, color-coded by secondary parameters

(2) Heatmaps: Parameter pair projections with aggregated performance metrics

(3) Time series: Population dynamics with confidence intervals across replicates

(4) Pareto frontiers: Multiobjective identification of optimal parameter regimes

Output directory structure.

```
wrappers_output/
+, vanilla/
| +, mba/          # MBA-only CSV files
| +, ba/           # BA-only CSV files
| +, delta.csv     # Delta analysis
| +, plots/        # Visualizations
| +, logs/         # Execution logs
| +, manifest.json # Complete metadata
```

```
+, stress/
| +, eps_0.01/     # Per-epsilon results
| +, eps_0.10/
| +, grid_summary.csv
+, grid_sweep/
| +, grid_summary.csv # All params
| +, plots/
| +, eda/           # Exploratory data analysis
+, continuous_sweep/
| +, continuous_summary.csv # Dense sampling
| +, experiments/   # Individual experiment specs
| +, perm_* /        # Per-permutation organization
```

H. Data collection and analysis

1. Fitness tracking

Daily fitness was calculated as the mean fitness across all five environmental states within each day. Final daily fitness for MBAs included subtraction of plasticity costs.

2. Population entropy

Shannon entropy was calculated from the distribution of phenotype sequences within each population: $H = -\sum[p_i \times \log_2(p_i)]$, where p_i is the frequency of sequence i .

3. Müller plots

Population genetic dynamics were visualized using Müller plots showing the frequency trajectories of the nine most common genotype sequences over time, with remaining sequences aggregated as “Other.”

4. Statistical analysis

All experiments used multiple independent replicates with results reported as mean $\pm$ standard error of the mean (SEM). Population-level dynamics were averaged across replicates, and evolutionary trajectories were analyzed using time-series methods.

I. Implementation details

The simulation was implemented in Python using NumPy for numerical computations and SciPy for statistical functions. The complete codebase is structured as a modular package with separate components for:

- (1) Agent classes (mba_vs_ba_sim.agents)
- (2) Environmental dynamics (mba_vs_ba_sim.env)
- (3) Population evolution (mba_vs_ba_sim.population)
- (4) Experimental scripts (scripts/)

All random number generation used NumPy’s PCG64 generator with explicit seed management for reproducibility. Computational experiments were designed to be fully deterministic given input parameters and random seeds.

J. Smart incubator platform

The Smart Incubator is an automated platform for long-duration experiments that impose temporally structured thermal stresses and sensory cues on microbial cultures. The system enables tests of phenotypic plasticity, prediction, and memory by controlling temperature and one or two cue modalities (LED light, vibration) under predefined temporal relationships. Full firmware, wiring diagrams, and bill of materials are available in the archived project repository [60] and are sufficient to reproduce the hardware and firmware implementation.

1. Hardware/firmware overview

The controller is an ESP32-based unit that drives a resistive heater and a Peltier cooler and reads a PT100 RTD via a MAX31865 interface. An OLED provides local status; a microSD card stores logs. The firmware (MicroPython) separates (i) hardware drivers, (ii) closed-loop temperature control, (iii) experiment logic (cycle scheduling and cue-heat correlation modes), and (iv) logging/safety. All pin-level details, PID constants, and watchdog settings are in the GitHub docs and are not required to interpret the experiments.

2. Thermal and cue control

Temperature is maintained in closed loop with a PID controller tuned for 15–40 °C operation. Heating and cooling actuators are driven by PWM and never coactivate. Typical heat shocks were to 32 °C (30 min) from a 23 °C baseline, with ramps of 2–3 min to heat and 5–8 min to cool. Cues were delivered by either a white LED (intensity specified as a duty cycle) or a miniature vibration motor (pulsed), as indicated in each figure legend.

3. Correlation modes

We implemented two cue-heat relationships: (i) *nonpredictive* (Mode 0), where cue timing within the basal phase is randomized independently of the heat-shock onset and (ii) *predictive* (Mode 1), where the cue precedes heat shock by a fixed interval (typically 30 min). Cycle durations were randomly sampled within user-specified bounds to minimize temporal entrainment. Exact bounds and the cue-heat interval used in each experiment are provided in the figure legends and archived repository documentation.

4. Logging and integrity

The device writes a snapshot of the system state at fixed intervals (10 s unless stated otherwise), including timestamp, measured/target temperature, controller output, cue states, and cycle/phase identifiers. A per-experiment manifest with SHA-256 hashes is generated for all files. The JSON schema, directory structure, integrity checks, and raw fields used in the analyses reported here are documented in the archived repository documentation.

5. Safety

Firmware enforces temperature limits and actuator duty-cycle caps and enters a safe state on sensor faults or SD errors.

Emergency stop can be triggered locally or via the optional web interface.

K. Experimental protocol framework

Experiments consisted of repeated basal and heat-shock phases organized into cycles with randomized durations, with default ranges described in the archived repository documentation. In predictive (Mode 1) runs, a cue was presented a fixed interval before heat shock; in nonpredictive (Mode 0) runs, cue timing was independent of heat shock. All default values appearing in figures correspond to those actually used for the specific dataset; no unused defaults are reported in the main text.

L. Data handling and availability

All device logs (JSON/CSV) and analysis notebooks are archived with the manuscript. Code and firmware are available in the archived project repository [60]. Analyses rely only on the fields defined in the archived repository documentation; any additional telemetry present in logs is not required to reproduce the results.

M. *Capsaspora* conditioning experiment

1. Culture conditions

Capsaspora owczarzaki (ATCC 30864) was cultured in ATCC medium 1034 at 23 °C. For conditioning experiments, cultures were grown to mid-log phase and distributed into Smart Incubator chambers.

2. Conditioning protocols

Two conditioning protocols were run in parallel over approximately one week (~35 cycles per condition):

(1) Predictive (Corr = 1): A compound conditioned stimulus (CS; LED light at 50% duty cycle + mechanical vibration, 30 s duration) was delivered immediately before each heat stress episode (32 °C for 30 min). CS-heat shock gaps ranged from 0.3 to 1.6 min.

(2) Nonpredictive (Corr = 0): The same CS was delivered at random times within each basal phase, independent of heat shock timing. CS-heat shock gaps ranged from −245 to +184 min.

Cycle durations were randomized (range: 240–400 min) to prevent temporal entrainment. Basal temperature was 23 °C; sublethal heat shocks were 32 °C for 30 min. To distinguish genuine habituation from sensory fatigue or metabolic depletion, our protocols follow the criteria established by Eckert *et al.* [61] for biochemically plausible single-cell learning: true habituation predicts (i) stimulus-specific response decrement, (ii) spontaneous recovery after rest, and (iii) dishabituation upon novel stimulus delivery. Fatigue, by contrast, predicts generalized response reduction and slow recovery dependent on metabolic state. In our model framework, habituation corresponds to decreased regret sensitivity (stable environments reduce exploration), while fatigue corresponds to reduced baseline fitness capacity.

3. Lethal challenge test

Following conditioning, all cultures received a priming cue (full-strength CS at basal temperature for 30 min), then

a lethal heat challenge (38 °C for 150 min). Cultures were immediately fixed and stained with propidium iodide (PI; 1 µg/ml for 15 min).

4. Image acquisition and analysis

Cultures were imaged on a Zeiss inverted microscope with a 20× objective. Fields were acquired tiling across the chamber. Cell segmentation was performed using Cellpose [62] with the “cyto2” model. For each segmented cell, mean PI fluorescence intensity was computed from the PI channel. Cells with PI mean $> 2.5 \times 10^5$ (arbitrary units) were classified as dead; the threshold was determined from the bimodal distribution of PI intensities across all cells.

5. Statistical analysis

Dead-cell fractions were computed per image. The difference between $\text{Corr} = 1$ and $\text{Corr} = 0$ conditions was assessed using a two-sided Mann-Whitney U test. Effect size was computed as the relative reduction in dead-cell fraction: $(\text{Corr}0 - \text{Corr}1)/\text{Corr}0$.

N. Information relegated to supplementary/repository documentation

Complete simulation source code (Python package implementing MBA/BA agents, population dynamics, and exper-

imental wrappers), hardware fabrication files (3D-printable STL models, PCB Gerber files, and KiCad schematics), firmware implementations (MicroPython for ESP32 controllers), parts lists, wiring diagrams, pin-level GPIO tables, PID tuning parameters, MQTT topic schemas, calibration notebooks, and example JSON telemetry are provided in the archived project repository [60]. These implementation details are not necessary to evaluate the theoretical framework or interpret the biological experiments when the cited repository release is available.

ACKNOWLEDGMENTS

We are grateful to Ricard Solé for his guidance, help with figure preparation, and thoughtful advice on the study’s methodology and overall scope. We also thank Roberto Paoli (PRBB uFabLab) for technical and conceptual support. This work was supported by grant PID2023-153273NB-I00 funded by MICIU/AEI/10.13039/501100011033 and FEDER, UE. We also acknowledge support from the Departament de Recerca i Universitats de la Generalitat de Catalunya (exp. 2021 SGR 00751).

DATA AVAILABILITY

The code, firmware, simulation scripts, and analysis materials supporting the findings of this article are openly available on Figshare [60].

- [1] R. G. Shaw and J. R. Etterson, Rapid climate change and the rate of adaptation: Insight from experimental quantitative genetics, *New Phytol.* **195**, 752 (2012).
- [2] P. W. Messer, S. P. Ellner, and N. G. Hairston, Can population genetics adapt to rapid evolution? *Trends Genet.* **32**, 408 (2016).
- [3] P. W. Messer and D. A. Petrov, Population genomics of rapid adaptation by soft selective sweeps, *Trends Ecol. Evol.* **28**, 659 (2013).
- [4] W. J. Blake, M. Kærn, C. R. Cantor, and J. J. Collins, Noise in eukaryotic gene expression, *Nature (London)* **422**, 633 (2003).
- [5] M. A. Coomer, L. Ham, and M. P. H. Stumpf, Noise distorts the epigenetic landscape and shapes cell-fate decisions, *Cell Syst.* **13**, 83 (2022).
- [6] P. Tataru, M. Simonsen, T. Bataillon, and A. Hobolth, Statistical inference in the Wright–Fisher model using allele frequency data, *Syst. Biol.* **66**, e30 (2017).
- [7] M. Kimura, Model of effectively neutral mutations in which selective constraint is incorporated, *Proc. Natl. Acad. Sci. USA* **76**, 3440 (1979).
- [8] R. Solé, Revisiting Leigh Van Valen’s “A new evolutionary law” (1973), *Biol. Theory* **17**, 120 (2022).
- [9] A. I. Khibnik and A. S. Kondrashov, Three mechanisms of Red Queen dynamics, *Proc. R. Soc. Lond. B* **264**, 1049 (1997).
- [10] L. Van Valen, A new evolutionary law, *Evol. Theory* **1**, 1 (1973).
- [11] M. M. Desai and D. S. Fisher, Beneficial mutation selection balance and the effect of linkage on positive selection, *Genetics* **176**, 1759 (2007).
- [12] B. H. Good, M. J. McDonald, J. E. Barrick, R. E. Lenski, and M. M. Desai, The dynamics of molecular evolution over 60,000 generations, *Nature (London)* **551**, 45 (2017).
- [13] S. F. Levy, J. R. Blundell, S. Venkataram, D. A. Petrov, D. S. Fisher, and G. Sherlock, Quantitative evolutionary dynamics using high-resolution lineage tracking, *Nature (London)* **519**, 181 (2015).
- [14] A. Forsman, Rethinking phenotypic plasticity and its consequences for individuals, populations and species, *Heredity* **115**, 276 (2015).
- [15] P. Lyon, F. Keijzer, D. Arendt, and M. Levin, Reframing cognition: Getting down to biological basics, *Phil. Trans. R. Soc. B* **376**, 20190750 (2021).
- [16] J. Zheng, J. L. Payne, and A. Wagner, Cryptic genetic variation accelerates evolution by opening access to diverse adaptive peaks, *Science* **365**, 347 (2019).
- [17] A. B. Paaby and M. V. Rockman, Cryptic genetic variation: Evolution’s hidden substrate, *Nat. Rev. Genet.* **15**, 247 (2014).
- [18] J. F. Ayroles, S. M. Buchanan, C. O’Leary, K. Skutt-Kakaria, J. K. Grenier, A. G. Clark, D. L. Hartl, and B. L. de Bivort, Behavioral idiosyncrasy reveals genetic control of phenotypic variability, *Proc. Natl. Acad. Sci. USA* **112**, 6706 (2015).
- [19] A. Sebé-Pedrós, M. Irimia, J. del Campo, H. Parra-Acero, C. Russ, C. Nusbaum, B. J. Blencowe, and I. Ruiz-Trillo, Regulated aggregative multicellularity in a close unicellular relative of Metazoa, *eLife* **2**, e01287 (2013).
- [20] N. Ros-Rocher, R. Q. Kidner, C. Gerdt, W. S. Davidson, I. Ruiz-Trillo, and J. P. Gerdt, Chemical factors induce aggregative

- multicellularity in a close unicellular relative of animals, *Proc. Natl. Acad. Sci. USA* **120**, e2216668120 (2023).
- [21] J. A. Schwartzman, A. Ebrahimi, G. Chadwick, Y. Sato, B. R. K. Roller, V. J. Orphan, and O. X. Cordero, Bacterial growth in multicellular aggregates leads to the emergence of complex life cycles, *Curr. Biol.* **32**, 3059 (2022).
- [22] R. B. Stelkens, M. A. Brockhurst, G. D. D. Hurst, and D. Greig, Hybridization facilitates evolutionary rescue, *Evol. Appl.* **7**, 1209 (2014).
- [23] M. J. West-Eberhard, *Developmental Plasticity and Evolution* (Oxford University Press, New York, 2003).
- [24] E. Jablonka and M. J. Lamb, *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*, revised ed. (MIT Press, Cambridge, MA, 2014).
- [25] M. Pigliucci and G. B. Müller, *Evolution: The Extended Synthesis* (MIT Press, Cambridge, MA, 2010).
- [26] K. N. Laland, T. Uller, M. W. Feldman, K. Sterelny, G. B. Müller, A. Moczek, E. Jablonka, and J. Odling-Smee, The extended evolutionary synthesis: Its structure, assumptions and predictions, *Proc. R. Soc. B* **282**, 20151019 (2015).
- [27] G. B. Müller, Why an extended evolutionary synthesis is necessary, *Interface Focus* **7**, 20170015 (2017).
- [28] D. Noble, A theory of biological relativity: No privileged level of causation, *Interface Focus* **2**, 55 (2012).
- [29] A. C. Alarcon Rios, L. B. Zwep, M. Vriesendorp, C. A. J. Knibbe, O. Ciofu, A. Liakopoulos, L. B. S. Aulin, D. E. Rozen, and J. G. C. van Hasselt, Colistin resistance dynamics in *Pseudomonas aeruginosa* under biofilm and planktonic growth, *Antimicrob. Agents Chemother.* **69**, e00421 (2025).
- [30] I. Dagogo-Jack and A. T. Shaw, Tumour heterogeneity and resistance to cancer therapies, *Nature Rev. Clin. Oncol.* **15**, 81 (2018).
- [31] G. S. França, M. Baron, B. R. King, J. P. Bossowski, A. Bjornberg, M. Pour, A. Rao, A. S. Patel, S. Misirlioglu, D. Barkley, K. H. Tang, I. Dolgalev, D. A. Liberman, G. Avital, F. Kuperwaser, M. Chiodin, D. A. Levine, T. Papagiannakopoulos, A. Marusyk, T. Lionnet, *et al.*, Cellular adaptation to cancer therapy along a resistance continuum, *Nature (London)* **631**, 876 (2024).
- [32] A. Oruni, M. Tchouakui, C. S. D. Tagne, J. Hearn, J. Kayondo, and C. S. Wondji, Temporal evolution of insecticide resistance and bionomics in *Anopheles funestus*, a key malaria vector in Uganda, *Sci. Rep.* **14**, 32027 (2024).
- [33] M. Acar, J. T. Mettetal, and A. van Oudenaarden, Stochastic switching as a survival strategy in fluctuating environments, *Nat. Genet.* **40**, 471 (2008).
- [34] H. I. Schreier, Y. Soen, and N. Brenner, Exploratory adaptation in large random networks, *Nat. Commun.* **8**, 14826 (2017).
- [35] C. H. Waddington, Genetic assimilation of an acquired character, *Evolution* **7**, 118 (1953).
- [36] E. Crispo, The Baldwin effect and genetic assimilation: Revisiting two mechanisms of evolutionary change mediated by phenotypic plasticity, *Evolution* **61**, 2469 (2007).
- [37] G. E. Hinton and S. J. Nowlan, How learning can guide evolution, *Complex Syst.* **1**, 495 (1987).
- [38] D. H. Ackley and M. L. Littman, Interactions between learning and evolution, in *Artificial Life II. SFI Studies in the Sciences of Complexity*, edited by C. G. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen (Addison-Wesley, Redwood City, CA, 1991), pp. 487–509.
- [39] K. Friston, The free-energy principle: A unified brain theory? *Nat. Rev. Neurosci.* **11**, 127 (2010).
- [40] D. V. Tikhonenkov, E. Hehenberger, A. S. Esaulov, O. I. Belyakova, Y. A. Mazei, A. P. Mylnikov, and P. J. Keeling, Insights into the origin of metazoan multicellularity from predatory unicellular relatives of animals, *BMC Biol.* **18**, 39 (2020).
- [41] N. Ros-Rocher, A. Pérez-Posada, M. M. Leger, and I. Ruiz-Trillo, The origin of animals: An ancestral reconstruction of the unicellular-to-multicellular transition, *Open Biol.* **11**, 200359 (2021).
- [42] J. E. Phillips, M. Santos, M. Konchwala, C. Xing, and D. Pan, Genome editing in the unicellular holozoan *Capsaspora owczarzaki* suggests a premetazoan role for the Hippo pathway in multicellular morphogenesis, *eLife* **11**, e77598 (2022).
- [43] A. Sebé-Pedrós, A. J. Roger, F. B. Lang, N. King, and I. Ruiz-Trillo, Ancient origin of the integrin-mediated adhesion and signaling machinery, *Proc. Natl. Acad. Sci. USA* **107**, 10142 (2010).
- [44] H. Suga, Z. Chen, A. de Mendoza, A. Sebé-Pedrós, M. W. Brown, E. Kramer, M. Carr, P. Kerner, M. Vervoort, N. Sánchez-Pons, G. Torruella, R. Derelle, G. Manning, B. F. Lang, C. Russ, B. J. Haas, A. J. Roger, C. Nusbaum, and I. Ruiz-Trillo, The *Capsaspora* genome reveals a complex unicellular prehistory of animals, *Nat. Commun.* **4**, 2325 (2013).
- [45] I. Ruiz-Trillo, K. Kin, and E. Casacuberta, The origin of metazoan multicellularity: A potential microbial black swan event, *Annu. Rev. Microbiol.* **77**, 499 (2023).
- [46] H. S. Jennings, *Behavior of the Lower Organisms* (Columbia University Press, New York, 1906).
- [47] P. Lyon, The cognitive cell: Bacterial behavior reconsidered, *Front. Microbiol.* **6**, 264 (2015).
- [48] F. Baluška, W. B. Miller, and A. S. Reber, Cellular and organismal agency—Not micromanagement by genes, *New Phytol.* **232**, 1337 (2021).
- [49] M. Levin, Collective intelligence of morphogenesis as a teleonomic process (2023), <https://doi.org/10.31234/osf.io/5ywqu>.
- [50] J. Gunawardena, Learning outside the brain: Integrating cognitive science and systems biology, *Proc. IEEE* **110**, 590 (2022).
- [51] J. Pla-Mauri and R. Solé, Engineering basal cognition: Minimal genetic circuits for habituation, sensitization, and massed-spaced learning, *ACS Synth. Biol.* **15**, 716 (2026).
- [52] C. I. Abramson and M. Levin, Behaviorist approaches to investigating memory and learning: A primer for synthetic biology and bioengineering, *Commun. Integr. Biol.* **14**, 230 (2021).
- [53] D. Noble, It's time to admit that genes are not the blueprint for life, *Nature (London)* **626**, 254 (2024).
- [54] K. Donohue, R. Rubio de Casas, L. Burghardt, K. Kovach, and C. G. Willis, Germination, postgermination adaptation, and species ecological ranges, *Annu. Rev. Ecol. Evol. Syst.* **41**, 293 (2010).
- [55] S. A. Frank and I. Yanai, The origin of novel traits in cancer, *Trends Cancer* **10**, 880 (2024).
- [56] A. Mitchell, G. H. Romano, B. Groisman, A. Yona, E. Dekel, M. Kupiec, O. Dahan, and Y. Pilpel, Adaptive prediction of environmental changes by microorganisms, *Nature (London)* **460**, 220 (2009).
- [57] S. D. Jamadar, A. Behler, H. Deery, and M. Breakspear, The metabolic costs of cognition, *Trends Cognit. Sci.* **29**, 541 (2025).

- [58] C. M. Jakobson and D. F. Jarosz, Metabolites control stress granule disassembly, *Nat. Cell Biol.* **23**, 1053 (2021).
- [59] M. Knafo, S. Rezenman, R. Nevo, I. Tsigalnikski, Z. Reich, and R. Kapon, Variability in protein expression and fitness under stress: Distinct modes of early adaptation of populations, *PRX Life* **3**, 013006 (2025).
- [60] M. Knafo, E. Casacuberta, and I. Ruiz-Trillo, Code, firmware, simulation scripts, and analysis materials for “Beyond diffusion: Bayesian learning strategies in single-cell life.” Figshare (2026), <https://doi.org/10.6084/m9.figshare.32235240>.
- [61] L. Eckert, M. S. Vidal-Saez, Z. Zhao, J. Garcia-Ojalvo, R. Martinez-Corral, and J. Gunawardena, Biochemically plausible models of habituation for single-cell learning, *Curr. Biol.* **34**, 5646 (2024).
- [62] C. Stringer, T. Wang, M. Michaelos, and M. Pachitariu, Cellpose: A generalist algorithm for cellular segmentation, *Nat. Methods* **18**, 100 (2021).