



The efficiency of information gain by natural selection

Ryan Seamus McGee^{a,1} , Olivia Kosterlitz^b , Artem Kaznatcheev^{c,d}, Benjamin Kerr^e , and Carl T. Bergstrom^{e,f}

Affiliations are included on p. 12.

Edited by Marcus W. Feldman, Stanford University, Stanford, CA; received April 21, 2026; accepted September 28, 2026

Natural selection enriches populations for genotypes that are well adapted to their environments, accumulating information about the selective conditions in the process. Therefore, selection can be viewed as a learning process through which populations acquire information about the environment. Motoo Kimura derived a relationship between the amount of information gained by selection and a measure of the inefficiency of the selective process, known as substitution load [M. Kimura, *Genet. Res.* 2, 127–140 (1961)]. Due to the gradual nature of selection and the ongoing mismatch of types to the environment, a lineage that is still acquiring information has lower mean fitness than a lineage that already has it. While this has been an influential insight, we find that Kimura's model is inconsistent with selection experiments involving *Escherichia coli*. To address limitations in this model and develop a more robust measure for the efficiency of selection, we turn to computational learning theory. We show that the learning-theoretic concept of regret quantifies the inefficiency of selection as a learning process in terms of relative lineage fitness. We establish bounds on the efficiency of information gain in general settings, including frequency-dependent selection and arbitrary environmental change. We experimentally validate these bounds and demonstrate that our extended theory recovers Kimura's fundamental insight. Finally, we show that selection is an asymptotically optimal algorithm for the basic problem faced by evolving populations and that it acquires information more efficiently than comparable processes in complex environments.

natural selection | information theory | learning theory | evolutionary game theory

To adapt to their environments, organisms require information about the environmental conditions and how to function in response. A central insight of evolutionary theory, dating back to Darwin, is that the process of natural selection acting on genetic variation can acquire adaptive information.

Selection changes type frequencies according to differences in reproductive success in the prevailing conditions. Over generations, these environment-dependent changes imprint the selective conditions on the population's composition, and relative type frequencies come to carry information about the environmental history. In this sense, selection can be viewed as a learning process that accumulates adaptive information. Several lines of work have drawn on information theory and learning theory to characterize what selection encodes and how rapidly that information can accumulate (1–11).

As selection updates the composition of the population to better match the environment, information about the environment is acquired, and the mean fitness of the population increases. But selection operates incrementally over time, and mean fitness remains suboptimal until a fully optimal composition is reached. The demographic consequences of the gradual nature of this process were first formalized by Haldane, who measured the “cost of selection” in terms of the reproductive excess required for a fitter allele to replace a less fit one (12, 13). Building on these ideas, Kimura showed that the amount of information gained during an allele substitution is directly proportional to the amount of potential lineage growth that goes unrealized because selection proceeds gradually (1). Kimura termed this quantity the substitution load, a growth-based formulation of Haldane's cost of selection.

Haldane and Kimura's “cost” framing has been criticized on the grounds that selection for beneficial mutations does not reduce fitness, and lineages that evolve are better off than those that do not (14–20). Instead, their cost of selection is a counterfactual one: It does not refer to the disadvantage of a lineage undergoing selection relative to one that does not evolve, but rather, it is the disadvantage of a lineage having to evolve relative to one that is already adapted. As Felsenstein (21) puts it, the cost of selection “comes from the fact that natural selection is less efficient than divine intervention.” While the

Significance

Natural selection changes the genetic makeup of a population in ways that provide information about its environmental history. In the 1960s, Motoo Kimura noted that a lineage that acquires information through selection has suboptimal fitness while the process unfolds, a seeming inefficiency arising from the gradual nature of selection. We analyze selection using computational learning theory, which allows us to formalize the efficiency of selection and establish general bounds on the efficiency of information gain by selection. Our extensions of this theory are consistent with experimental populations and reaffirm mathematical relationships between information and fitness. We show that fitness “regret” bounds information gain but that selection nevertheless acquires information more efficiently than comparable processes in variable environments.

Author contributions: R.S.M., B.K., and C.T.B. designed research; R.S.M. and O.K. performed research; R.S.M. and O.K. analyzed data; O.K. and A.K. provided feedback on drafts; B.K. and C.T.B. supervised research and provided feedback on drafts; and R.S.M. and A.K. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2026 the Author(s). Published by PNAS. This article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

¹To whom correspondence may be addressed. Email: ryansmcgee@gmail.com.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2611968123/-DCSupplemental>.

Published October 29, 2026.

outcome of positive selection is beneficial, there may be more or less efficient processes by which those improvements are made.

Kimura's result belongs to a broader class of limits on how rapidly selection can accumulate information. Worden measured information gain as the concentration of an initially uniform population on favored types and showed that this entropy reduction cannot exceed cumulative logarithmic selection pressure (2). Worden's result recovers Kimura's one-for-one information-load relationship for asexual populations and generalizes the constraint to a "speed limit" on long-run information gain that accommodates both sexual reproduction and time-varying fitnesses (2).

Subsequent work extends this perspective to finite stochastic populations subject to mutation and drift (3, 5, 8, 11). In this setting, Mustonen and Lässig (8) and Hledik et al. (11) quantify information attributable to selection relative to neutral evolution and bound its accumulation using population-size-scaled measures of fitness flux and parental-sampling KL divergence, respectively, emphasizing how information gain is governed by variation in reproductive success.

Learning-theoretic approaches ask the complementary question of how effectively evolution uses available fitness differences. Bayesian interpretations of selection cast relative reproductive success as evidence that reweights the alternatives present in a population (22–24). Valiant's theory of evolvability concerns whether repeated local variation and selection can discover high-performing representations within feasible evolutionary time (25). Chastain and colleagues made the connection to cumulative performance explicit by showing that weak selection at independently recombining loci implements Multiplicative Weights Updating, which balances cumulative fitness payoff against the retention of entropy (10, 26–29).

Together, this literature constrains information gain by selection in two complementary senses. Fitness differences limit how strongly selection can distinguish among alternatives, while evolutionary dynamics determine how effectively those differences are integrated over time. How these constraints relate to cumulative lineage fitness, the quantity central to Haldane and Kimura, remains less clear. We address this question by comparing an evolving lineage with a fixed reference composition exposed to the same sequence of environments, separating the type-condition mismatch that is unavoidable by any fixed composition from the additional shortfall associated with the gradual nature of selection.

We begin by formalizing selection as an information acquisition process and reviewing Kimura's result relating information gain to substitution load. We then test this result in laboratory experiments, which prompts us to examine how load should be measured in heterogeneous and time-varying environments. This leads us to identify *regret*, a measure of relative lineage fitness, as the proper metric for evaluating the demographic consequences of selection as a learning process in general settings. Finally, we relate these lineage fitness quantities to information, deriving general bounds showing that information gain cannot exceed regret. In doing so, we establish guarantees on the efficiency of selection, including under frequency-dependent selection and arbitrary environmental change.

Natural Selection as an Information Acquisition Process

Learning is the process of iteratively updating a hypothesis in light of new evidence. In the context of evolution, consider a lineage

defined as a population of asexual replicators of alternative types (e.g., genotypes) descended from a shared ancestral population. Each type represents a strategy for survival and reproduction, and the type-frequency distribution $\mathbf{p}^t$ can be viewed as the lineage's hypothesis about which type is best suited to the environment at time t . The fitness W_i of type i (i.e., its expected reproductive output) provides evidence about its suitability for the current environment, and natural selection updates the frequency p_i of each type according to standard haploid replicator dynamics (*SI Appendix, Appendix A.1.1*):

$$p_i^{t+1} = \frac{p_i^t W_i}{\langle W \rangle_{\mathbf{p}^t}}, \quad [1]$$

where $\langle W \rangle_{\mathbf{p}^t} = \sum_i p_i^t W_i$ is the mean fitness of the lineage and its one-generation growth factor. Throughout, growth quantities refer to expected growth before density regulation or other type-independent demographic effects and therefore need not correspond to realized changes in population size.

Selection updates type frequencies according to relative fitness, just as Bayesian learning adjusts hypothesis weights according to the likelihood of the observed evidence under each hypothesis (22–24, 30, 31) (*SI Appendix, Appendix C.1*). We can measure how much a lineage learns by measuring how much its hypothesis changes. The Kullback–Leibler (KL) divergence between the lineage's initial type distribution $\mathbf{p}^0$ and its updated distribution $\mathbf{p}^T$ quantifies the amount of *information gain* I^T that selection provides over a learning period of duration T (Fig. 1A):

$$I^T := D(\mathbf{p}^T \| \mathbf{p}^0) = \sum_i p_i^T \log \frac{p_i^T}{p_i^0}. \quad [2]$$

The information gained by a lineage is "about" the environmental history that induced the corresponding frequency changes, in the sense that experiencing that history constrains the range of type compositions likely to be observed. For example, when the initial type distribution is maximally unconstrained (uniform over types), information gain equals the selection-induced reduction in entropy of the type distribution, a quantity referred to as the "specific information" of the population (4, 32). More generally, when $\mathbf{p}^0$ is nonuniform, I^T retains an analogous interpretation as the average reduction in type surprisal from the initial to the postselection distribution, weighted by endpoint frequencies (*SI Appendix, Appendix B.2.1*).

Across lineages, these environment-dependent frequency changes establish an association between types and conditions such that observing a type reduces uncertainty about the selective history in which its lineage evolved, and vice versa (33–35). This association is quantified by the mutual information $MI^T(G; E)$, where E denotes an environmental history and G the type of an individual sampled at time T from the corresponding lineage. For descendant lineages that share a common ancestral type distribution $\mathbf{p}^0$, let $\mathbf{p}_e^T$ and I_e^T denote the endpoint type distribution and information gain under each lineage's history e . The mutual information accumulated by time T is the average information gain across histories minus a correction for frequency changes shared across histories (35) (*SI Appendix, Appendix B.2.3*):

$$MI^T(G; E) = \langle I_e^T \rangle_E - D(\langle \mathbf{p}_e^T \rangle_E \| \mathbf{p}^0). \quad [3]$$

I^T measures the information gained by one lineage along a realized selective history. Across lineages, these gains contribute

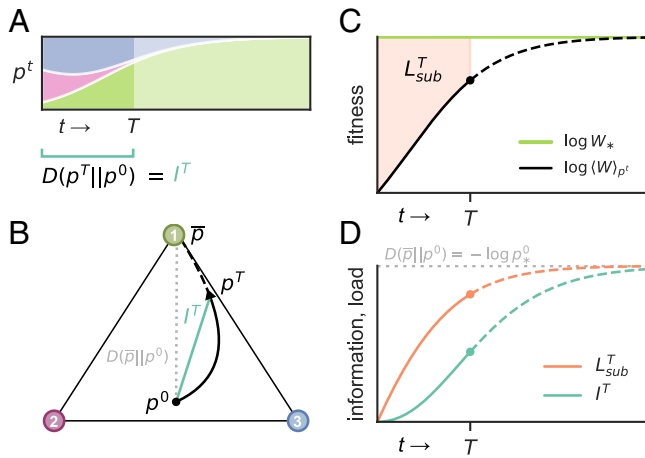


Fig. 1. Selection accumulates information gain and substitution load. (A) Changes in the population's type frequencies over time are depicted by a Muller plot, with the height of each colored band giving the frequency of the corresponding type [type colors are indicated in (B)]. Information gain I^T is the KL divergence of the population's composition $\mathbf{p}^T$ from its initial composition $\mathbf{p}^0$. (B) The black trajectory tracks the population's composition as selection moves it from its initial distribution $\mathbf{p}^0$ toward fixation of the fittest Type 1 (green vertex), which is a stationary learning target $\bar{\mathbf{p}}$. The teal segment from $\mathbf{p}^0$ to $\mathbf{p}^T$ provides graphical intuition for the information gained by time T (although KL divergence is not a distance metric). The gray dotted segment depicts the initial potential information $D(\bar{\mathbf{p}} \parallel \mathbf{p}^0)$. (C) As selection proceeds in this fixed environment, the population's per-generation growth rate $\log \langle W \rangle_{\mathbf{p}^t}$ increases (black line) and approaches the fittest-type growth rate $\log W_*$ (green line). The orange shaded region represents the substitution load accumulated through time T , the cumulative growth-rate shortfall relative to a population fixed for Type 1 from the outset. (D) Information gain (teal line) and substitution load (orange line) accumulate over the course of selection. Substitution load is at least as large as information gain at every time (Proposition 2). At fixation, both converge to the initial potential information $D(\bar{\mathbf{p}} \parallel \mathbf{p}^0) = -\log p_*^0$, where p_*^0 is the initial frequency of Type 1 (Proposition 1).

to the accumulation of mutual information between types and environments, which quantifies the portion of these gains that distinguishes among alternative histories (35) (SI Appendix, Appendix B.2.3). By contrast, frequency changes due to drift, mutation, or other processes whose probabilities are independent of environmental history do not increase this mutual information in expectation, although they may modify an association already established by selection (11, 35) (SI Appendix, Appendix B.2.3). For simplicity, we focus our current analysis on infinite populations of haploid, asexual replicators where selection is the only process acting on type frequencies.

Information gain tracks selection's reweighting of a lineage toward better-suited types, which is accompanied by an increase in mean fitness (6, 36, 37). Over an allele substitution, the information gained bounds the accompanying increase in lineage growth rate (4, 32, 38) (SI Appendix, Appendix B.2.1):

$$\lim_{T \rightarrow \infty} I^T \geq \log W_* - \log \langle W \rangle_{\mathbf{p}^0}. \quad [4]$$

Here $\log W_*$ is the per-generation growth rate of the lineage following fixation of the fittest type, and $\log \langle W \rangle_{\mathbf{p}^0}$ is its initial per-generation growth rate (SI Appendix, Appendix B.2.1 gives the corresponding finite-time relationship). Thus selection does not shift a lineage to a higher fitness composition without gaining a corresponding amount of information.

In some settings, selection moves a lineage's composition toward a fixed composition $\bar{\mathbf{p}}$. The divergence $D(\bar{\mathbf{p}} \parallel \mathbf{p}^t)$ of that target from the lineage's current type distribution defines the **potential information**: The information remaining to be

acquired as the lineage approaches it (Fig. 1B). When selection converges to $\bar{\mathbf{p}}$ while progressively reducing this potential information, $\bar{\mathbf{p}}$ serves as a *learning target*. An accessible evolutionarily stable state is one important example of such a target, although an attracting target need not be evolutionarily stable (7, 39–41). As the lineage converges to a learning target, potential information approaches zero and information gain approaches the initial potential information (Fig. 1D and SI Appendix, Appendix B.2.2).

Information Gain Is Associated with Load

In the simple case of an allele substitution, the fittest type increases in frequency toward fixation, but the continued presence of less-fit types keeps the population's mean growth rate below the optimum until the substitution is complete. Kimura's *substitution load* L_{sub}^T measures the resulting cumulative growth shortfall over T generations relative to a counterfactual population fixed for the fittest type from the outset (1) (Fig. 1C and SI Appendix, Appendix D.1.1):

$$L_{\text{sub}}^T := \sum_{t=0}^{T-1} \left(\log W_* - \log \langle W \rangle_{\mathbf{p}^t} \right), \quad [5]$$

where $\log W_*$ is the per-generation growth rate of the fittest type and $\log \langle W \rangle_{\mathbf{p}^t}$ is the per-generation growth rate of the evolving lineage with composition $\mathbf{p}^t$.

Effectively, substitution load captures the cumulative growth disadvantage of reaching an optimal composition gradually rather than instantaneously. Kimura showed that substitution load converges to $-\log p_*^0$ as the fittest type approaches fixation, where p_*^0 is its initial frequency.

Kimura quantified information gain based on the ratio of the fixation probability under selection to that under drift alone (1 and p_*^0 , respectively). The resulting quantity, $-\log p_*^0$, coincides with the KL divergence measure of information gain I^T in the limit of fixation. Accordingly, in Kimura's model, information gain and substitution load converge to this common value (Fig. 1D):

Proposition 1 (1).

$$\lim_{T \rightarrow \infty} I^T = \lim_{T \rightarrow \infty} L_{\text{sub}}^T = -\log p_*^0. \quad [6]$$

Worden later obtained an equivalent limiting relationship for a population with a uniform initial type distribution and also showed that substitution load remains at least as large as information gain throughout the substitution (2) (Fig. 1D). We extend this finite-time bound to arbitrary initial compositions:

Proposition 2 (SI Appendix, Appendix D.4.2).

$$I^T \leq L_{\text{sub}}^T \quad \forall T. \quad [7]$$

This suggests that gaining a given amount of information is associated with a minimum substitution load. In other words, acquiring one bit of information, which represents a two-fold reduction in uncertainty about the environment, requires at least a two-fold shortfall in cumulative growth relative to a population that grows optimally from the beginning.

Measuring Load in Variable Environments

These results predict a relationship between two primary quantities in evolutionary biology: fitness and information. While this theory is clear and self-consistent, it was developed under several simplifying assumptions and has not been directly validated. It is important to confront such theory with the real world to assess its applicability to actual populations and uncover opportunities to accommodate more complex systems.

We set out to test these predictions using an experimental system that was designed to closely adhere to the theory's assumptions. We used four strains of *Escherichia coli* with distinct mutations in the RNA polymerase *rpoB* gene, each conferring a different growth rate in a standard growth medium. We performed selection competitions with various strain combinations and initial frequencies, each of which proceeded for 36 h as a sequence of four-hour serial-transfer growth intervals (*Materials and Methods* and *SI Appendix, Appendix E.2.1*). Information gain was measured at each interval as the KL divergence between the population's current and initial allele frequency distributions. Interval growth rates were estimated from corresponding density measurements and transfer volumes (*SI Appendix, Appendix E.2*). Load was quantified as the cumulative difference between the population's average growth rate and the growth rate of the fittest allele in each interval.

Results from one competition are shown in detail in *Fig. 2 B–F*. As expected, the fastest-growing allele increased in frequency and approached fixation (*Fig. 2B*). However, the substitution load did not always exceed information gain or converge on information gain at fixation as predicted (*Fig. 2D*). This is because Kimura's definition of substitution load assumes a static

environment with constant fitnesses, but in our experiments, growth rates fluctuated stochastically despite maintaining cultures in the exponential growth phase (*Fig. 2C*). If this assumption cannot be satisfied even in highly controlled laboratory conditions, applying this fixed definition of load to natural systems is problematic.

These observations led us to extend Kimura's concept of load to contexts where conditions and fitnesses change over time. Suppose there is a set of distinct environmental conditions characterized by abiotic properties, type frequencies, population densities, or other factors (*SI Appendix, Appendix A.2*). Let x_j^t denote the probability that a given individual experiences condition j in generation t , and let W_{ij} denote the fitness of type i in condition j . Then scenarios with fluctuating conditions can be represented by a distribution $\mathbf{x}^t$ that changes arbitrarily over time.

To extend the comparison underlying substitution load to variable environments, we benchmark the lineage against the growth attainable under ideal type–condition matching, in which each condition is paired with the type best suited to it. If W_{*j} is the fitness of the fittest type in condition j , then **mismatch load** is the cumulative shortfall in population growth rate relative to this ideal benchmark (*SI Appendix, Appendix D.1.2*):

$$L^T := \sum_{t=0}^{T-1} \left(\log \langle W_* \rangle_{\mathbf{x}^t} - \log \langle W \rangle_{\mathbf{p}^t, \mathbf{x}^t} \right). \quad [8]$$

Here, $\log \langle W_* \rangle_{\mathbf{x}^t}$ is the growth rate under ideal type–condition matching, and $\log \langle W \rangle_{\mathbf{p}^t, \mathbf{x}^t}$ is the growth rate of the evolving lineage, averaged over its type composition and the conditions experienced. Like substitution load, mismatch load quantifies the

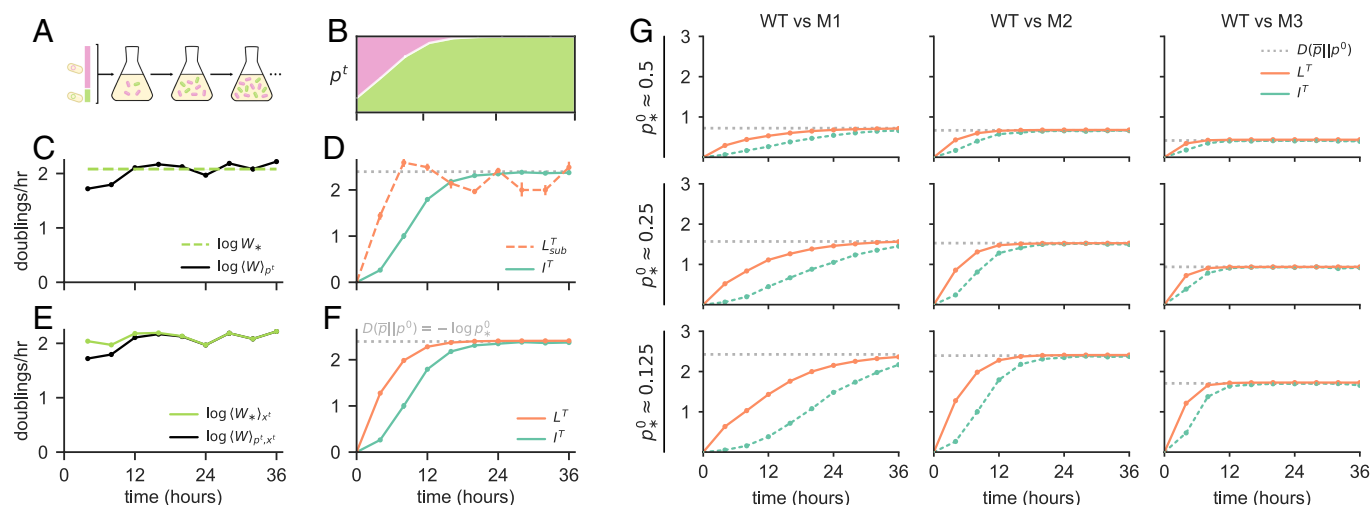


Fig. 2. The predicted relationship between mismatch load and information gain is observed across selection experiments. (A) Selection competitions paired a GFP-labeled WT strain with a less-fit RFP-labeled mutant strain. (B–F) Results from one representative WT–M2 competition are shown in detail. (B) The WT and M2 strains began at frequencies of approximately 12.5% ($p_*^0 = 0.125$) and 87.5%, respectively. Their frequencies over time are depicted by a Muller plot. Selection moves the population toward fixation of the fittest WT strain. (C) The black series gives the population growth rate during each preceding 4-h interval. The green dashed line gives a fixed estimate of the characteristic WT growth rate, calculated as the average of the interval-specific WT estimates. (D) Information gain (teal) is compared with *substitution* load (dashed orange), calculated as the cumulative difference between the growth rates in (C). Because growth rates fluctuate, this load does not satisfy the predictions of *Proposition 2* or *Proposition 1*. (E) The black series repeats the population growth rates from (C), while the green series gives the interval-specific WT growth rates. The population growth rate approaches the WT rate as the WT strain nears fixation. (F) *Mismatch* load (orange), calculated using the interval-specific WT rates in (E), is compared with information gain (teal). Mismatch load remains at least as large as information gain, and both approach the predicted common value $D(\bar{\mathbf{p}}||\mathbf{p}^0) = -\log p_*^0$. (G) Results from nine competitions span three WT–mutant pairs (columns) and three approximate initial WT frequencies (rows). The WT strain was fittest and approached fixation in every competition. Gray dotted lines show the initial potential information $D(\bar{\mathbf{p}}||\mathbf{p}^0) = -\log p_*^0$, the predicted common limiting value of mismatch load and information gain. Mismatch load is at least as large as information gain throughout each competition, and both approach this limit. The limiting value depends on the initial WT frequency rather than the strength of selection, which differs among mutants. Throughout (B–G), points and lines show mean estimates, and vertical error bars show corresponding 95% measurement intervals propagated from replicate cytometry enumerations (*SI Appendix, Appendix E.2.2*); these intervals are often smaller than the point markers.

potential lineage growth foregone because types are mismatched with environmental conditions over time. When the same type has the highest fitness in every condition, a population fixed for that type attains the ideal-matching benchmark. When different conditions favor different types, ideal-matching may not be achievable by any feasible fixed composition.

Despite efforts to maintain a static and identical environment in our experiments, detailed conditions (e.g., inoculum size, nutrient concentration, temperature) nevertheless varied subtly from one interval to the next. This caused growth rates to fluctuate, but the identity of the fittest type did not change, so this remained an allele substitution scenario in which the theoretical relationship between information gain and load was expected to hold (*SI Appendix, Appendix D.4*). Using the mismatch load framework, the empirical results align well with the theoretical predictions. Mismatch load was at least as large as information gain, and the two quantities approached equality as populations neared fixation (*Fig. 2F*). This relationship was observed across all selection competitions, spanning different selection strengths and initial frequencies of the fittest allele (*Fig. 2G* and *SI Appendix, Fig. S8*). Therefore, our extended model and experimental results recover Kimura's original insight: there is a minimum amount of load associated with every bit of information gain by natural selection.

A Learning-Theoretic Framework for Evaluating the Efficiency of Selection

Analyzing the relative performance of learning processes is the purview of computational learning theory, which has seen recent adoption for the study of evolution (*10, 25, 42*). Here we use this framework to clarify the meaning of mismatch load and develop a rigorous, general measure of the efficiency of selection as a learning process.

The learning problem faced by an evolving population can be modeled as repeated play of a game between two players: the population and the environment (*Fig. 3A* and *SI Appendix, Appendix C.2*). In each round t (i.e., generation t) of this *Population versus Environment* (PvE) game, the population 'plays' a distribution of types $\mathbf{p}^t$, and the environment 'plays' a distribution of conditions $\mathbf{x}^t$. Each individual experiences a condition independently drawn from $\mathbf{x}^t$, so the resulting type-condition pairings are random and outside the control of either player (*Fig. 3B*). A specific learning problem is defined by the fixed fitness W_{ij} of each type i in each condition j and the process that generates the environmental sequence $\mathbf{x}^0, \mathbf{x}^1, \dots, \mathbf{x}^T$. Appropriate choices of these fitnesses and environmental distributions can represent a wide range of biological settings, including environments that are heterogeneous across individuals and variable over time.

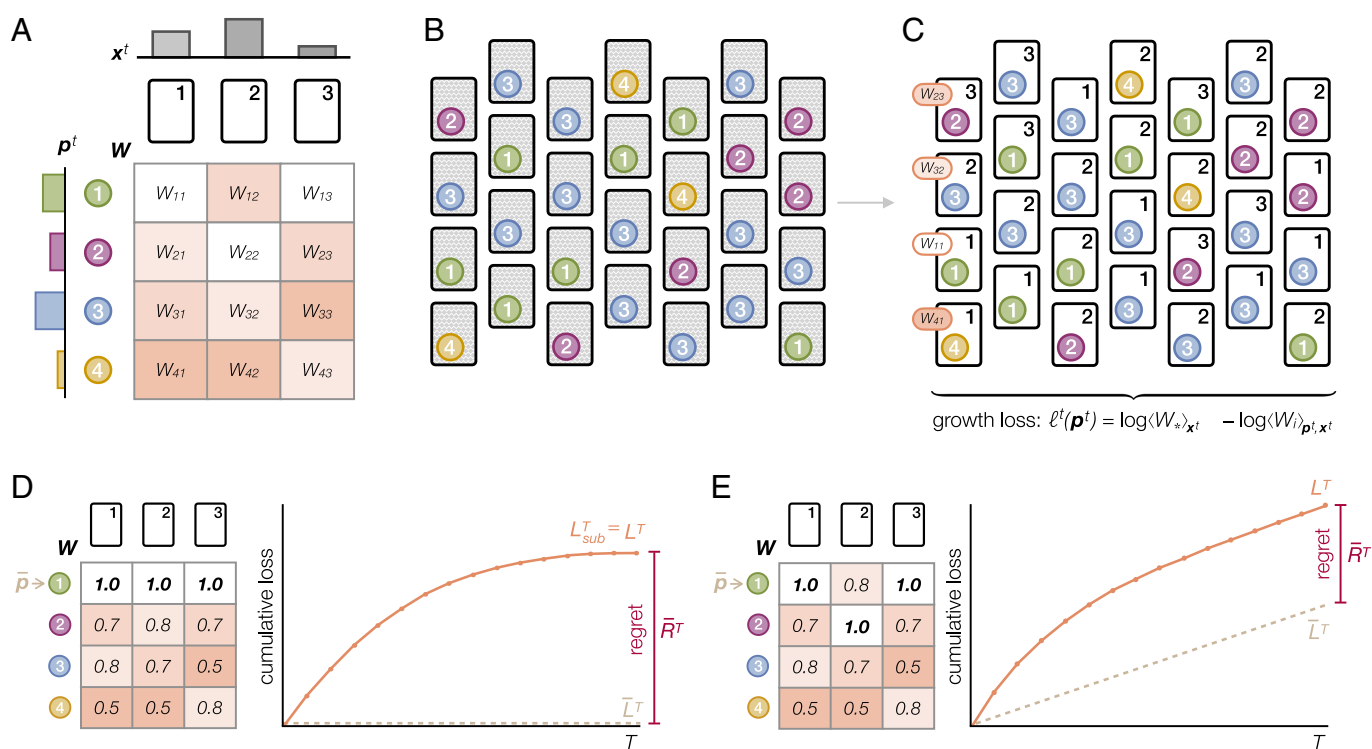


Fig. 3. Regret isolates the inefficiency of learning in the Population versus Environment game. The learning problem faced by an evolving population can be modeled as a repeated game against the environment, which is illustrated using the metaphor of a tabletop game. (A) The game is defined by the set of types (tokens, rows), the set of environmental conditions (cards, columns), and a game matrix $\mathbf{W}$ that assigns fitness W_{ij} to each type i in each condition j . (B) In round t , the population and environment play the frequency distributions $\mathbf{p}^t$ and $\mathbf{x}^t$, respectively. Each individual is assigned to an independent environmental condition (one token placed on each card), but the pairing of types and conditions is random (cards are shuffled and remain face down until tokens are placed). (C) The revealed type-condition pairings determine the fitness W_{ij} of each individual (selected examples annotated). The population's one-round growth loss $\ell^t(\mathbf{p}^t)$ is the shortfall in its per-generation growth rate relative to ideal type-condition matching. The population updates $\mathbf{p}^t$ from this fitness feedback, and its performance is evaluated by cumulative growth loss. (D) Type 1 is fittest in every condition, so fixation of Type 1 is a stationary target $\bar{\mathbf{p}}$ with zero mismatch load ($\bar{L}^T = 0$). In such a case, mismatch load reduces to substitution load ($L^T_{\text{sub}} = L^T$), and target regret coincides with the population's total load ($\bar{R}^T = L^T - \bar{L}^T = L^T$, red bar). (E) Different conditions favor different types, so no fixed composition can attain ideal type-condition matching. Therefore, the stationary target $\bar{\mathbf{p}}$ (fixation of Type 1) accumulates baseline mismatch load $\bar{L}^T$ (gold dashed line). The population's mismatch load L^T (orange curve) includes this baseline, while target regret $\bar{R}^T = L^T - \bar{L}^T$ isolates the excess load attributable to the learning process (red bar).

The fitnesses experienced in each round provide feedback about the relative suitability of the available types, which the lineage uses to update its type distribution. We evaluate the performance of the lineage's overall composition using its one-round *growth loss*: the shortfall in the lineage's growth rate relative to ideal type-condition matching (Fig. 3C):

$$\ell^t(\mathbf{p}^t) := \log \langle W_* \rangle_{\mathbf{x}^t} - \log \langle W \rangle_{\mathbf{p}^t, \mathbf{x}^t}. \quad [9]$$

The cumulative population growth loss over T generations equals the mismatch load L^T defined in Eq. 8.

Mismatch Load Does Not Always Reflect the Inefficiency of Selection. Substitution load was originally conceived as a “cost of selection” arising from the gradual nature of the process (1, 12, 13, 43–45). When there is only one environmental condition (as in Haldane and Kimura's substitution models) or when the same type has the highest fitness in every condition (as in our experiments), a population fixed for that type attains the ideal-matching benchmark and experiences no mismatch load (Fig. 3D). Any mismatch is therefore transient and attributable to the substitution being incomplete, and mismatch load directly quantifies the growth disadvantage of reaching the optimal composition gradually through selection.

When different conditions favor different types, however, their random association with individuals may cause even the best feasible fixed composition to accumulate a baseline amount of mismatch load (Fig. 3E). As a result, total mismatch load reflects this baseline in addition to any inefficiency of the evolutionary process itself. To isolate the latter, we compare the evolving lineage's load with that of a fixed reference composition exposed to the same environmental sequence. This difference, formalized below as regret, isolates the excess load associated with the evolutionary trajectory.

Regret Formalizes the Inefficiency of Selection. In learning theory, *regret* is the difference between the cumulative loss incurred by a learner as it updates its strategy over time and the loss it could have achieved had it played a fixed reference strategy from the beginning (Fig. 3 D and E). Regret thereby formalizes the inefficiency of a learning process: the excess loss one experiences in using that process to learn an effective strategy as opposed to knowing that solution all along.

For an evolving population, the appropriate fixed reference depends on the sequence of environments and the biological question being asked. We first consider settings in which selection approaches a fixed learning target $\bar{\mathbf{p}}$. If the learning target accessible to the population remains the same throughout the sequence of environments, then selection will continually approach and ultimately settle on this *stationary learning target* (SI Appendix, Appendix C.3). Such a target may be polymorphic, as can occur when fluctuating selection maintains genetic variation (46). Its composition is determined by selection's convergence dynamics and provides a natural reference for measuring the growth shortfall incurred by selection's incremental approach. We later consider a standard learning-theoretic hindsight comparison, which accommodates more complex learning problems where selection is not necessarily directed toward a single stationary target.

As a lineage gradually converges on the target, it accumulates mismatch load. Its *target regret* $\bar{R}^T$ relative to the stationary learning target $\bar{\mathbf{p}}$ is the difference between its load L^T and the load $\bar{L}^T$ of a counterfactual lineage that maintains the target

composition throughout the same environmental sequence:

$$\bar{R}^T := L^T - \bar{L}^T \quad [10]$$

$$= \sum_{t=0}^{T-1} \left(\log \langle W \rangle_{\bar{\mathbf{p}}, \mathbf{x}^t} - \log \langle W \rangle_{\mathbf{p}^t, \mathbf{x}^t} \right). \quad [11]$$

This quantity can equivalently be expressed as the negative log ratio of the evolving lineage's cumulative growth factor Γ^T to that of the counterfactual target lineage $\bar{\Gamma}^T$: $\bar{R}^T = -\log(\Gamma^T / \bar{\Gamma}^T)$ (SI Appendix, Appendix D.2.2). We can interpret the ratio $\Gamma^T / \bar{\Gamma}^T$ as the *relative lineage fitness*: the cumulative growth of the evolving lineage relative to the counterfactual lineage maintained at the target composition.

The lineage's “goal” in the game can be interpreted as learning the target composition as efficiently as possible so as to maximize its relative lineage fitness. Because target regret is inversely related to relative lineage fitness, maximizing lineage fitness is equivalent to minimizing regret. For a fixed target composition and environmental sequence, minimizing regret is in turn equivalent to minimizing mismatch load. All of these objectives involve keeping the evolving lineage's cumulative growth as close as possible to that of the counterfactual lineage that maintains the target composition throughout. Regret is a useful metric not because it yields a different optimization, but because it isolates the shortfall attributable to the evolutionary trajectory.

Natural Selection as a No-Regret Learning Process

In general, it is intuitive for the population player to adapt its strategy by increasing the weight of types that have incurred low growth loss in the past while maintaining diversity to hedge against future conditions. A simple but powerful learning algorithm known as *Multiplicative Weights Updating* (MWU) balances this tradeoff and has been shown to be effective in settings where the environment can change arbitrarily over time (47–49) (SI Appendix, Appendix C.2.3). A learning-rate parameter controls the balance between concentrating weight on types that have performed well and maintaining diversity. With bounded losses and an appropriately tuned learning rate, MWU's average performance provably approaches that of the best fixed strategy in hindsight, and its average per-round regret vanishes in the long run, a property known as *no-regret* learning.

Chastain and others have shown that weak-selection allele-frequency dynamics at multiple segregating loci implement MWU (10, 27–29). In the asexual setting considered here, applying MWU to the type-level growth loss $\ell^t(i) = \log \langle W_* \rangle_{\mathbf{x}^t} - \log \langle W_i \rangle_{\mathbf{x}^t}$ yields the replicator update. The population growth loss $\ell^t(\mathbf{p}^t)$ (Eq. 9) is the corresponding MWU *mix loss*, which measures the growth shortfall of the population as a whole. With growth losses measured using natural logarithms, replicator dynamics implement MWU with a fixed learning rate of one, which gives equal weight to minimizing loss and maintaining diversity (SI Appendix, Appendix C.2.4). However, because selection's learning rate is fixed rather than tuned to the learning problem, standard performance guarantees for MWU do not carry over automatically. Moreover, both standard MWU guarantees and Chastain et al.'s selection-specific applications evaluate externally averaged loss over types rather than mix loss, so neither establishes bounds on regret in terms of the actual cumulative growth of the

evolving lineage (10, 26, 27) (*SI Appendix, Appendixes C.2.4 and C.2.5*).

So far, we have shown that information gain provides a tight lower bound on growth regret in the special case of an allele substitution, where substitution load, mismatch load, and target regret coincide. We now characterize selection's efficiency more broadly by bounding population growth regret in more general evolutionary settings.

When the sequence of environments is characterized by a stationary learning target $\bar{\mathbf{p}}$, the total mismatch load accumulated by a lineage undergoing selection is bounded.

Proposition 3 (*SI Appendix, Appendix D.3.2*).

$$L^T \leq \bar{L}^T + D(\bar{\mathbf{p}} \parallel \mathbf{p}^0) \quad \forall T. \quad [12]$$

This bound distinguishes two sources of cumulative mismatch: the initial potential information, which reflects the divergence of the lineage's initial state from the target, and the mismatch load that the target itself experiences across the realized environmental sequence.

Accordingly, the initial potential information is an upper bound on regret with respect to a stationary learning target.

Theorem 1 (*SI Appendix, Appendix D.3.2*). *For a lineage beginning at $\mathbf{p}^0$ with an accessible stationary learning target $\bar{\mathbf{p}}$, target regret is bounded by the initial potential information*

$$\bar{R}^T \leq D(\bar{\mathbf{p}} \parallel \mathbf{p}^0) \quad \forall T, \quad [13]$$

with convergence from below: $\lim_{T \rightarrow \infty} \bar{R}^T = D(\bar{\mathbf{p}} \parallel \mathbf{p}^0)$.

Theorem 1 shows that total target regret is finite, ensuring that the average regret approaches zero in the long run (i.e., $\bar{R}^T / T \rightarrow 0$ as $T \rightarrow \infty$). Therefore, selection is a no-regret algorithm with respect to a stationary learning target and population growth loss. This result complements no-regret results associated with Chastain et al.'s external loss formulations (10, 26, 27). In this sense, natural selection is an asymptotically optimal algorithm for solving the fundamental learning problem faced by evolving populations.

These bounds are exact in continuous time and in the weak-selection limit of discrete-time replicator dynamics. In general, discrete updates introduce a nonnegative Jensen correction, resulting from the difference between averaging the per-generation growth rates of types and taking the growth rate of their mixture (*SI Appendix, Appendix D.3.1*). This correction vanishes in the weak-selection limit or whenever all types represented in the target have equal effective fitness (such as fixation of a single type; *SI Appendix, Appendix D.3.2*). For simplicity, main-text results involving a stationary learning target are stated for these zero-correction cases; The *SI Appendix* gives the fully general discrete-time results with the correction included.

The Efficiency of Information Acquisition by Selection

Kimura's allele substitution setting is a natural starting point for evaluating how efficiently selection acquires information. Throughout a substitution, cumulative load is at least as large as information gain (*Proposition 2* and *Fig. 2*), and the two converge to the same value at fixation (*Proposition 1*). Having established regret as a more general measure of efficiency, we can extend this comparison beyond allele substitutions and quantify the regret selection incurs per unit of information gained.

The connection between information gain and lineage fitness is apparent at the scale of a single selective update. In discrete time, the information acquired in a single generation, $D(\mathbf{p}^{t+1} \parallel \mathbf{p}^t)$, cannot exceed the accompanying improvement in lineage growth rate with respect to the current environment:

Proposition 4 (*SI Appendix, Appendix D.4.1*).

$$\begin{aligned} D(\mathbf{p}^{t+1} \parallel \mathbf{p}^t) &\leq \ell^t(\mathbf{p}^t) - \ell^t(\mathbf{p}^{t+1}) \\ &= \log \langle W \rangle_{\mathbf{p}^{t+1}, \mathbf{x}^t} - \log \langle W \rangle_{\mathbf{p}^t, \mathbf{x}^t}. \end{aligned} \quad [14]$$

We can view this loss difference as a one-step "look-ahead regret" that the lineage has about not already having the composition enjoyed by the next generation. If this regret were zero, the lineage would have neither fitness nor information to gain in updating its composition. Thus, the experience of regret must precede the acquisition of information.

As a lineage converges to an accessible stationary learning target $\bar{\mathbf{p}}$, its cumulative information gain approaches the initial potential information, acquiring the total amount of information that can be gained for the given target (*SI Appendix, Appendix D.4.2*). Together, this convergence and the preceding target regret bound establish a common limiting value for information gain and target regret.

Theorem 2 (*SI Appendix, Appendix D.4.2*). *For a lineage beginning at $\mathbf{p}^0$ with an accessible stationary learning target $\bar{\mathbf{p}}$, information gain and target regret converge, satisfying*

$$\lim_{T \rightarrow \infty} I^T = \lim_{T \rightarrow \infty} \bar{R}^T = D(\bar{\mathbf{p}} \parallel \mathbf{p}^0). \quad [15]$$

This result extends Kimura's one-for-one relationship beyond allele substitutions to convergence on any stationary target. Each fold reduction in uncertainty comes with a corresponding fold reduction in lineage fitness relative to a lineage that had the target composition from the outset. When no stationary target exists, information gain remains bounded by growth regret relative to the best fixed composition in hindsight, as developed below for arbitrary environmental change.

Complex Learning Problems for Evolving Populations

Frequency Dependence. Within the PvE framework, the environmental distribution need not be constant or exogenous. In fact, there are no restrictions on how the environment's distribution of conditions can change over time. For example, the environment's strategy in each round may depend on the population's current or prior compositions, which allows for cases of frequency-dependent selection. In particular, if the frequencies of conditions experienced by individuals are set by the population's current composition (i.e., $\mathbf{x}^t = \mathbf{p}^t$), then the population defines its own environment and the PvE model reduces to a conventional autonomous evolutionary game. This setting opens up additional dynamics, including social dilemmas, and can transform the adaptive capabilities of evolving populations (42, 50).

The frequency-dependent setting highlights why selection's learning is evaluated through *relative* lineage fitness (i.e., regret) rather than fitness in absolute terms. To illustrate this, consider the Prisoner's Dilemma described in *SI Appendix, Fig. S2*, a scenario in which $\mathbf{x}^t = \mathbf{p}^t$. The best collective outcome occurs if the population plays an all-cooperate strategy, which achieves the

maximum possible population mean fitness. However, replicator dynamics lead toward an all-defect population, which continually reduces mean fitness along the way.

The only feedback available to the population player is the set of effective fitnesses (losses) of each type in each round. At each step, the defector type has higher effective fitness (lower loss) than the cooperator, and replicator dynamics increase the frequency of defectors accordingly. Arriving at an all-cooperate strategy would require the population player to use information about the game or its frequency dependence that is not available from the observed fitnesses alone or to adopt a learning algorithm that is not generally robust across learning problems. Thus the population may simply be unable to maximize absolute fitness using available feedback.

In general, the relevant learning-theoretic benchmark is the lineage fitness achieved by the best fixed composition against the same realized sequence of environments. This fixed comparison may seem strange when the environmental sequence itself depended on the population's trajectory while learning was underway, but interpreting regret in terms of relative lineage fitness lends intuition.

For example, consider a large, well-mixed population of cells that evolves according to replicator dynamics in the Prisoner's Dilemma scenario (SI Appendix, Fig. S2). Suppose we identify two subpopulations, each large enough for drift to be negligible but small relative to the overall population. The first subpopulation begins with the same type distribution as the overall population, and replicator dynamics increase the frequency of defectors, tracking the compositional change in the overall population. The second subpopulation is fixed for defectors from the beginning, which turns out to be optimal in hindsight. Because these subpopulations are small relative to the overall population, neither affects the population composition that constitutes the "environment." They therefore experience the same sequence of environments, and it is natural to ask how the cumulative growth of the two lineages compares over that shared history.

In the Prisoner's Dilemma, the mean fitness of both lineages decreases as the surrounding population shifts toward defection (SI Appendix, Fig. S2). However, the all-defect lineage has higher mean fitness than the learning lineage at every step. The learning lineage thus experiences an actual instantaneous growth-rate shortfall and achieves less cumulative growth over the selection period. Relative lineage fitness (regret) quantifies this comparison. Just as the growth of a learning lineage can be compared to a fixed subpopulation in the same physical environment, a learning process can be evaluated with respect to the best fixed composition in hindsight, whether or not such a comparator lineage actually existed.

Arbitrary Environmental Change. We now extend our analysis of growth regret to the fully general setting where no restrictions are placed on the sequence of environments. When the distribution of environmental conditions $\mathbf{x}^t$ is free to change arbitrarily, selection may be directed toward different compositions over time or may fail to converge altogether. Whereas an environmental sequence characterized by a stationary target presents a basic learning problem with a fixed solution, a sequence in which the target changes from one horizon to the next can be viewed as a series of learning problems across which the lineage must integrate its learning.

If we make no restrictions on the sequence of environments, then we cannot determine a fixed optimal composition based

on the process that generates the environmental sequence or on the dynamical trajectory of the population, since these are free to change at any time. Instead, we use the standard learning-theoretic benchmark of the best fixed composition in hindsight. At horizon T , the *hindsight-optimal composition* $\tilde{\mathbf{p}}^T$ is the fixed composition accessible from $\mathbf{p}^0$ that would have minimized cumulative population growth loss over the realized environmental sequence (SI Appendix, Appendix D.2.3).

We refer to regret measured with respect to the hindsight-optimal composition as *hindsight regret*:

$$\tilde{R}^T := L^T - \tilde{L}^T \quad [16]$$

$$= \sum_{t=0}^{T-1} \left(\log \langle W \rangle_{\tilde{\mathbf{p}}^T, \mathbf{x}^t} - \log \langle W \rangle_{\mathbf{p}^t, \mathbf{x}^t} \right), \quad [17]$$

where $\tilde{L}^T$ is the cumulative mismatch load of a counterfactual lineage that maintains $\tilde{\mathbf{p}}^T$ throughout the realized sequence of environments.

Stationary learning targets and hindsight-optimal compositions play different roles. A stationary target is distinguished by the convergence dynamics, whereas the hindsight-optimal composition is identified retrospectively from the realized sequence. They may coincide in some settings but need not do so in general. When environmental conditions favor different compositions at different times, $\tilde{\mathbf{p}}^T$ may change as the evaluation horizon grows. The examples in Fig. 4 B and C illustrate this distinction: the Rock–Paper–Scissors population approaches a stationary target even as its "environment" changes, whereas the hindsight-optimal composition in the cycling environment occasionally switches as the evaluation horizon grows.

Nevertheless, the lineage's hindsight regret is bounded in this arbitrary environmental change setting.

Theorem 3 (SI Appendix, Appendix D.3.3). *Under discrete-time replicator dynamics, hindsight regret is bounded:*

$$\tilde{R}^T \leq D(\tilde{\mathbf{p}}^T \| \mathbf{p}^0) + c\tilde{L}^T \quad \forall T \quad [18]$$

where $c := \ell_{\max}/(1 - e^{-\ell_{\max}}) - 1$ for type growth losses in the range $0 \leq \ell^t(i) \leq \ell_{\max}$. This bound imposes no restrictions on the fitness values or sequence of environments that constitute the learning problem. Therefore, this bound provides a general worst-case guarantee for the total hindsight regret a lineage can experience.

As the environmental distribution changes, compositions that perform well during one part of the sequence may perform poorly during another. This can contribute to the mismatch load that even the best fixed composition in hindsight accumulates. Because this load enters the regret bound through the term $c\tilde{L}^T$, the general discrete-time bound does not guarantee a finite upper limit on hindsight regret independent of T .

The excess regret term $c\tilde{L}^T$ bounds the cumulative Jensen correction introduced by finite discrete updates, which arises if the hindsight-optimal composition is mixed and its types have unequal effective fitness (SI Appendix, Appendix D.3.3). When all types in the hindsight-optimal composition have equal effective fitness, this correction vanishes and the bound reduces to $\tilde{R}^T \leq D(\tilde{\mathbf{p}}^T \| \mathbf{p}^0)$ (SI Appendix, Appendix D.3.3). This zero-correction form is also approached in the weak-selection limit and holds exactly under continuous-time replicator dynamics.

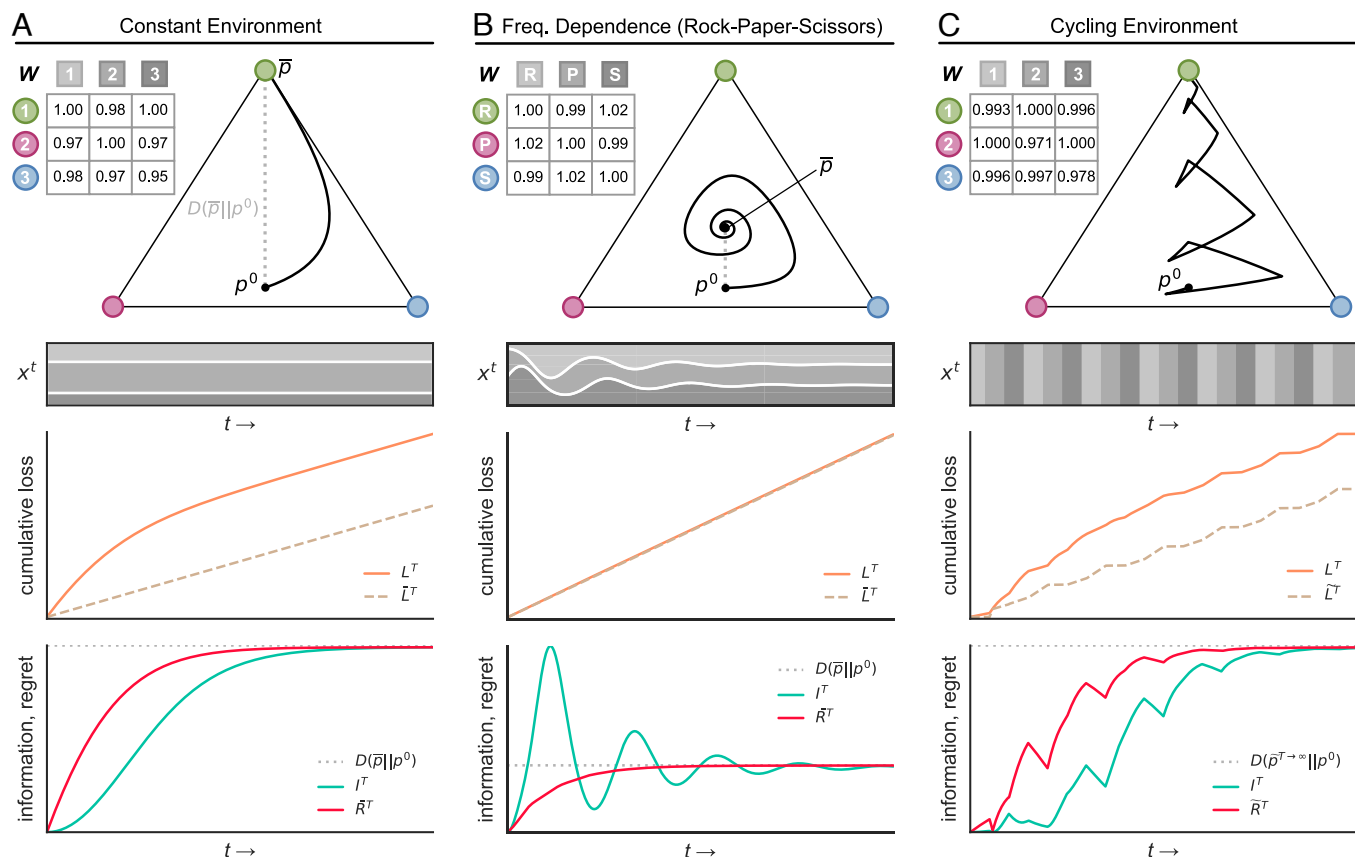


Fig. 4. Fundamental relationships between regret and information gain hold across environmental contexts. Simulations of replicator dynamics in three environmental contexts are shown in columns. In each column, the Population versus Environment game is characterized by a fixed matrix $\mathbf{W}$ that gives the fitness W_{ij} of type i in environmental condition j . Selection moves the population composition along the black trajectory in each simplex. The panels below the simplexes show the distribution of environmental conditions $\mathbf{x}^t$ over time, with each vertical slice representing the distribution in one generation. The *Middle* panels compare the mismatch load L^T of the evolving population with either the target load $\tilde{L}^T$ (columns A and B) or the load $\tilde{L}^T$ of the horizon-specific hindsight-optimal composition (column C). The *Bottom* panels show information gain I^T together with target regret $\tilde{R}^T$ (columns A and B) or hindsight regret $\tilde{R}^T$ (column C). Gray dotted references mark the initial divergence from $\mathbf{p}^0$ to the relevant limiting composition. (A) The distribution of environmental conditions is constant. Type 1 (green) has the greatest effective fitness and increases in frequency in every generation. Because Type 1 is not optimal in every condition, the stationary target consisting of Type 1 fixation accumulates mismatch load $\tilde{L}^T$ (dashed gold line). Target regret is bounded by the initial potential information $D(\bar{\mathbf{p}}||\mathbf{p}^0)$, and target regret and information gain converge to this value (*Theorems 1 and 2*). (B) The environmental conditions are the types themselves, with $\mathbf{x}^t = \mathbf{p}^t$ in every generation. In this setting, the PvE game defined by the matrix $\mathbf{W}$ coincides with a conventional autonomous evolutionary game. The fitness matrix defines a Rock-Paper-Scissors interaction in which each type is favored against one type and disfavored against another. Selection produces damped oscillations and converges to a mixed stationary target $\bar{\mathbf{p}}$. Information gain fluctuates before converging to the initial potential information $D(\bar{\mathbf{p}}||\mathbf{p}^0)$. Target regret approaches the same value from below (*Theorems 1 and 2*). (C) The fitness matrix $\mathbf{W}$ remains fixed, while the environmental distribution $\mathbf{x}^t$ cycles periodically among three pure conditions. The favored type varies by condition, so no stationary learning target exists. As the retrospective horizon T grows, the hindsight-optimal fixed composition $\bar{\mathbf{p}}^T$ changes before eventually settling on fixation of Type 1, which has the greatest long-run growth across the cycle. Selection also converges toward Type 1, so hindsight regret remains finite and approaches the initial divergence from the long-run hindsight-optimal composition $D(\bar{\mathbf{p}}^{T \rightarrow \infty}||\mathbf{p}^0)$ (*Theorem 3*). At every horizon, hindsight regret bounds information gain from above (*Theorem 4*).

This upper bound limits how far the evolving lineage can fall behind the best fixed composition. Information gain provides a complementary lower bound on the same growth shortfall. Unlike the information gained while approaching a stationary target, information gain in an arbitrarily changing environment may not increase monotonically. Because $I^T = D(\mathbf{p}^T||\mathbf{p}^0)$ measures the net difference between the lineage's current and initial compositions, later selective changes can reinforce or reverse changes induced by earlier conditions. These fluctuations can be interpreted as the lineage gaining information about the present environment and "unlearning" information it had acquired about conditions that are no longer relevant.

Even when information gain fluctuates due to environmental change, its relationship to relative lineage fitness persists at every horizon.

Theorem 4 (*SI Appendix, Appendix D.4.3*). *Under discrete-time replicator dynamics, information gain is bounded by hindsight regret:*

$$I^T \leq \tilde{R}^T \quad \forall T. \quad [19]$$

The information acquired through the lineage's compositional change cannot exceed its cumulative growth shortfall relative to the best fixed composition in hindsight. Regret may exceed information gain because it records this shortfall along the entire evolutionary trajectory, whereas information gain records only the net difference between the lineage's initial and current compositions.

Therefore, each unit of information requires at least one unit of regret, but the worst-case regret associated with information gain by natural selection is bounded. Together, these results confirm a

direct, general relationship between information gain and relative lineage fitness.

Comparing Selection with Alternative Learning Rates

The correspondence between selection and Multiplicative Weights Updating allows us to ask how the same update would perform if it responded more or less strongly to fitness differences. The responsiveness of this process is controlled by a learning rate η . Natural selection implements MWU with the fixed learning rate $\eta = 1$ (SI Appendix, Appendix C.2.4). Lower rates ($\eta < 1$) dampen the effect of fitness differences and reallocate the lineage more conservatively, whereas higher rates amplify their effect and reallocate it more aggressively (SI Appendix, Appendix D.3.4). Faster updating can move the lineage quickly toward a consistently successful composition but can also respond too strongly to transient fitness advantages. Slower updating reduces the risk of overreaction, but takes even longer than selection to approach a good composition.

Worst-case regret bounds provide a standard benchmark for comparing learning processes. Extending Theorem 3, we derive general bounds on hindsight growth regret across fixed learning rates (SI Appendix, Appendix D.3.4). The learning rate giving the tightest worst-case bound depends on two features of the learning problem. More aggressive updating is favored when the hindsight-optimal composition lies far from the initial lineage (large $D(\tilde{\mathbf{p}}^T \parallel \mathbf{p}^0)$), because more compositional change is required to reach it. By contrast, selection’s less aggressive response is favored when even the hindsight-optimal composition accumulates substantial load (large $\tilde{L}^T$).

Two clear results follow from these rate-dependent bounds (SI Appendix, Fig. S1). First, learning rates below that of selection always produce a worse regret bound. Slowing the update does not improve the performance of the hindsight-optimal composition; it only delays the lineage’s movement toward it. Second, rates above that of selection can improve the bound when the hindsight-optimal composition experiences little load, because a more responsive process approaches it more rapidly. However, this advantage of being more responsive than selection diminishes as the load of the hindsight-optimal composition increases.

Once hindsight-optimal load becomes sufficiently large relative to the initial potential information, selection’s learning

rate minimizes the regret bound over all fixed response rates (SI Appendix, Fig. S1). Persistent mismatch in heterogeneous or time-varying environments can raise the total hindsight-optimal load toward this selection-optimizing regime. Initial potential information is bounded, so if hindsight-optimal load continues to grow over time, selection’s learning rate will eventually minimize the bound among all fixed response rates. While this result concerns worst-case guarantees rather than every realized trajectory, it suggests that selection’s learning rate is well suited to the variable conditions that evolving populations face over long time horizons.

Beyond this fixed-type comparison, the noninstantaneous response of selection may also promote evolutionary exploration through adaptive momentum, in which mutations that initially reduce fitness remain in an expanding clade long enough for subsequent mutations to carry the lineage across a fitness valley (51). Further delineating the environments in which selection outperforms other learning processes remains an important direction for future research.

Discussion

Our results, summarized in Table 1, describe an intimate relationship between fitness and information. A learning-theoretic view of evolution identifies growth regret, a measure of relative lineage fitness, as a natural measure of inefficiency in the selection process. When the learning target is stationary, selection approaches that target and achieves vanishing per-generation regret. More generally, hindsight regret separates the growth shortfall associated with the lineage’s evolutionary trajectory from the mismatch that even the best fixed composition cannot avoid. We recover Kimura’s original allele-substitution relationship experimentally and show more generally that information gain cannot exceed regret. However, comparing regret bounds across fixed learning rates suggests that selection is particularly well suited for information acquisition in environments where persistent mismatch accumulates over long periods.

Chastain and colleagues previously established the learning-theoretic foundation of this perspective by showing that, under weak selection, allele-frequency dynamics at recombining loci implement MWU (10, 27). They applied standard fixed-comparator MWU guarantees to bound each locus’s cumulative fitness-payoff shortfall relative to the best single allele in hindsight, an external-loss comparison analogous to our hindsight growth-regret comparison. Chastain and Smith further showed

Table 1. Summary of results relating information gain and regret under increasingly general conditions (for zero-correction settings)

Environment	Homogeneous (1 condition)	Heterogeneous		
	Constant ($\mathbf{x}^t = \mathbf{x}^0 \forall t$)	Constant	Time-varying	
	Stationary target ($\bar{\mathbf{p}}$)	Stationary target	Stationary target	Arbitrary
Information bound	$I^T \xrightarrow{\gamma} L_{\text{sub}}^T$ Kimura (1), Worden (2) Proposition 2	$I^T \xrightarrow{T} \bar{R}^T$ Theorem 2	$I^T \xrightarrow{T} \bar{R}^T$ Theorem 2	$I^T \leq \tilde{R}^T \forall T$ Theorem 4
Regret bound	$L_{\text{sub}}^T \xrightarrow{\gamma} -\log p_*^0$ Kimura (1), Worden (2) Proposition 1	$\bar{R}^T \xrightarrow{\gamma} -\log p_*^0$ Theorem 1	$\bar{R}^T \xrightarrow{\gamma} D(\bar{\mathbf{p}} \mathbf{p}^0)$ Theorem 1	$\tilde{R}^T \leq D(\bar{\mathbf{p}}^T \mathbf{p}^0) + c\tilde{L}^T \forall T$ Theorem 3

The arrow notation $a^T \xrightarrow{\gamma} b$ signifies convergence ($\lim_{T \rightarrow \infty} a^T = b$), and the angled arrow notation $a^T \xrightarrow{\gamma} b$ signifies convergence from below ($\lim_{T \rightarrow \infty} a^T = b$ and $a^T \leq b \forall T$).

that, in i.i.d. environments, external no-regret learning over a finite set of strategies implies asymptotically matching the growth rate of the best fixed strategy (26). Both sets of guarantees are based on type-averaged external losses, rather than the cumulative demographic growth performance of the evolving mixed population as a whole. Our framework distinguishes population growth loss from external loss, makes the population growth comparison explicit, and derives bounds connecting cumulative mismatch load, relative lineage fitness, and initial compositional information. Most importantly, these bounds place general limits on selection's information gain in terms of its cumulative growth performance, with the hindsight results applying to arbitrary realized environmental sequences.

These results fit within a broader body of work relating selection, fitness, and information. Worden analyzed a similar infinite-population model with time-varying selection pressures (2). For a uniform initial composition, his measure of haploid genotype information coincides with our information gain, while his cumulative logarithmic selection pressure corresponds to mismatch load when fitness is measured relative to the best type in each generation. The argument underlying Worden's evolutionary "speed limit" therefore implicitly bounds information gain by total mismatch load in the special case of an allele substitution. This result recovers Kimura's limiting equality and also bounds information gain by load throughout the substitution. Our framework further extends this relationship to arbitrary learning problems and provides a tighter bound on information gain in terms of regret rather than total load.

Other work extends these relationships to stochastic finite populations, where information is measured by divergence between selected and neutral distributions over population states rather than by the displacement of a single lineage from its initial composition (3, 5, 8, 11). Of particular relevance to this work, Hledík et al. bound this ensemble information by a "cost of selection" quantity, defined as the cumulative KL bias that relative fitness introduces into parental sampling. Their cost per parental sample is exactly the divergence $D(\mathbf{p}^{t+1} \parallel \mathbf{p}^t)$ that we identify as one-step information gain. The identity underlying our *Proposition 4* decomposes the corresponding look-ahead growth-rate improvement into this information plus a Jensen gap, which gives their sampling cost bound a direct lineage-growth interpretation.

Hledík et al. and complementary work by Mustonen and Lässig highlight the role of fitness variation in constraining information gain, whereas our cumulative performance measures, like those of Kimura and Worden, are expressed primarily in terms of mean fitness (8, 11). However, our discrete-time bounds also account for this variation through Jensen gap terms whose leading term is quadratic in fitness differences. Although our presentation emphasizes limits where these terms vanish, the general bounds retain them and describe how fitness variation affects the relationship between information gain and growth regret. Applied to Hledík et al.'s per-sample cost, the single-step regret decomposition of *Proposition 4* recovers their weak-selection relationship between information and relative fitness variance and sharpens their general variance bound (*SI Appendix, Appendix D.4.1*).

The finite-population bounds of Hledík et al. and Mustonen and Lässig scale with population size and therefore allow more information gain per unit of fitness-based constraint in larger populations. This scaling reflects the potential for greater statistical distinction between selected and neutral ensembles in larger populations. Hledík et al.'s multilocus analysis further

highlights that recombination can reduce interference among selected loci and substantially increase the efficiency of selection relative to asexual reproduction (11). Our regret bounds instead isolate a population-size-independent relationship between a frequency trajectory and its cumulative performance in asexual populations. As a result, our bounds provide a basis for comparing the efficiency of selection with that of other learning processes in this setting, while extending these growth-regret bounds to recombining populations is an area for future work.

Our results concern deterministic selection in infinite haploid asexual populations with a fixed set of types, and therefore focus on acquiring information from standing variation rather than the broader problem of evolutionary search. In general, minimizing regret relative to the best currently available composition need not maximize long-term evolvability. Mutation expands the type set and moves populations across adaptive landscapes, where epistasis can create fitness valleys and make finding optima computationally hard (52). Recombination likewise changes which combinations are available to selection, which can aid traversal by joining beneficial variants but also disrupt favorable epistatic combinations. Extending comparator-based measures of lineage performance to processes that generate and reorganize variation may help reconcile selection's efficient exploitation of standing variation with the demands of evolutionary exploration.

Materials and Methods

Bacterial Strains. *Escherichia coli* B (REL606) strains were used in the selection experiments. A "wild type" strain (WT) and three strains with unique mutations in the RNA polymerase *rpoB* gene (M1, M2, M3) were obtained with permission from the strain archive of Lindsey et al. (53). Mutations to the *rpoB* gene gave each mutant strain a distinct exponential growth rate that was reduced from that of the WT strain (*SI Appendix, Fig. S3*). As the fastest-growing strain, the WT strain was transformed with a plasmid carrying the green fluorescent protein (GFP) gene *mGFPmut2*. Each *rpoB* mutant strain was transformed with a plasmid carrying the red fluorescent protein (RFP) gene *mScarlet-I*. A full description of all strains and plasmids used in this study can be found in *SI Appendix, Appendix E.1*.

Selection Experiments. Strains were cultured in standard Luria-Bertani (LB) broth with 15 µg/mL tetracycline for marker plasmid retention. Selection competition experiments were conducted between three pairs of strains: WT versus M1, WT versus M2, and WT versus M3. Each pair of strains participated in three competitions with approximate initial WT frequencies of 12.5%, 25%, and 50% of the initial population. Competitions were initiated by combining strains in the designated ratios at a density of $\sim 1 \times 10^6$ CFU/mL in 25 mL of fresh media. Every 4 h, a 1 mL sample was taken from each competition culture to measure strain densities and frequencies using flow cytometry (*SI Appendix, Appendix E.2.2*). At the same time, a sample of the culture was transferred to 25 mL of fresh media using an adaptive transfer protocol that ensured that the culture would remain in exponential growth while maintaining a measurable density for flow cytometry (*SI Appendix, Appendix E.2.1*). Competitions proceeded in this fashion for 36 h.

Data, Materials, and Software Availability. Data files are archived and available at <https://zenodo.org/records/10183554> (54). All other data are included in the manuscript and/or *SI Appendix*.

ACKNOWLEDGMENTS. We thank the referees for their thoughtful reviews, which strengthened the manuscript's framing and positioning relative to prior work. We are grateful to Phillip Wacker for his insightful feedback that prompted substantial improvements to our formulations and presentation. Throughout this project, we benefited from valuable discussions with Michael Lachmann, Kevin Gross, Frazer Meacham, Joe Felsenstein, Emily Brunelli, and

the Kerr Lab. We also thank Katie Dickinson for exceptional guidance with experimental techniques and logistics, and Reilly Falter and Alena Beloborodyy for their contributions at the bench. R.S.M. was supported by the NSF Graduate Research Fellowship Program (DGE-2140004) and by the NSF Bio/computational Evolution in Action Consortium (BEACON) Center. O.K. was supported by the NSF Graduate Research Fellowship Program (DGE-1762114) and by the NSF Postdoctoral Research Fellowships in Biology Program (DBI-2305907). A.K. was supported by a fellowship from the James S. McDonnell Foundation.

C.T.B. was supported in part by Templeton World Charity Foundation grant 20650.

Author affiliations: ^aDepartment of Neuroscience, Washington University School of Medicine in St. Louis, St. Louis, MO 63110; ^bDepartment of Biological Sciences, University of Idaho, Moscow, ID 83844; ^cDepartment of Mathematics, Utrecht University, Utrecht 3508 TA, The Netherlands; ^dDepartment of Information and Computing Sciences, Utrecht University, Utrecht 3584 CC, The Netherlands; ^eDepartment of Biology, University of Washington, Seattle, WA 98195-1800; and ^fSanta Fe Institute, Santa Fe, NM 87501

1. M. Kimura, Natural selection as the process of accumulating genetic information in adaptive evolution. *Genet. Res.* **2**, 127–140 (1961).
2. R. P. Worden, A speed limit for evolution. *J. Theor. Biol.* **176**, 137–152 (1995).
3. Y. Iwasa, Free fitness that always increases in evolution. *J. Theor. Biol.* **135**, 265–281 (1988).
4. C. Adami, C. Ofria, T. C. Collier, Evolution of biological complexity. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 4463–4468 (2000).
5. G. Sella, A. E. Hirsh, The application of statistical physics to evolutionary biology. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 9541–9546 (2005).
6. S. A. Frank, Natural selection maximizes Fisher information. *J. Evol. Biol.* **22**, 231–244 (2009).
7. M. Harper, Information geometry and evolutionary game theory. arXiv [Preprint] (2009). <https://arxiv.org/abs/0911.1383> (Accessed 3 September 2019).
8. V. Mustonen, M. Lässig, Fitness flux and ubiquity of adaptive evolution. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 4248–4253 (2010).
9. M. C. Donaldson-Matasci, C. T. Bergstrom, M. Lachmann, The fitness value of information. *Oikos* **119**, 219–230 (2010).
10. E. Chastain, A. Livnat, C. Papadimitriou, V. Vazirani, Algorithms, games, and evolution. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 10620–10623 (2014).
11. M. Hledik, N. Barton, G. Tkačik, Accumulation and maintenance of information in evolution. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2123152119 (2022).
12. J. B. S. Haldane, The effect of variation of fitness. *Am. Nat.* **71**, 337–349 (1937).
13. J. B. S. Haldane, The cost of natural selection. *J. Genet.* **55**, 511–524 (1957).
14. L. Van Valen, Haldane's dilemma, evolutionary rates, and heterosis. *Am. Nat.* **97**, 185–190 (1963).
15. A. M. Brues, The cost of evolution vs. the cost of not evolving. *Evolution* **18**, 379–383 (1964).
16. J. Maynard Smith, "Haldane's dilemma" and the rate of evolution. *Nature* **219**, 1114–1116 (1968).
17. J. A. Sved, Possible rates of gene substitution in evolution. *Am. Nat.* **102**, 283–293 (1968).
18. P. O'Donald, "Haldane's dilemma" and the rate of natural selection. *Nature* **221**, 815–816 (1969).
19. P. A. P. Moran, "Haldane's dilemma" and the rate of evolution. *Ann. Hum. Genet.* **33**, 245–249 (1970).
20. J. Matheson, M. Exposito-Alonso, J. Masel, Substitution load revisited: A high proportion of deaths can be selective. *Genetics* **229**, iyaf011 (2025).
21. J. Felsenstein, On the biological significance of the cost of gene substitution. *Am. Nat.* **105**, 1–11 (1971).
22. M. Harper, The replicator equation as an inference dynamic. arXiv [Preprint] (2009). <https://arxiv.org/abs/0911.1763> (Accessed 26 February 2019).
23. C. R. Shalizi, Dynamics of Bayesian updating with dependent data and misspecified models. *Electron. J. Stat.* **3**, 1039–1074 (2009).
24. R. A. Watson, E. Szathmáry, How can evolution learn? *Trends Ecol. Evol.* **31**, 147–157 (2016).
25. L. G. Valiant, Evolvability. *J. ACM* **56**, 1–21 (2009).
26. E. Chastain, C. Smith, Process information and evolution. *IEEE Trans. Mol. Biol. Multiscale Commun.* **2**, 240–248 (2016).
27. E. Chastain, "The multiplicative weight updates method for evolutionary biology," PhD thesis, Rutgers University, School of Graduate Studies, New Brunswick, NJ (2017).
28. R. Mehta, I. Panageas, G. Piliouras, Natural selection as an inhibitor of genetic diversity: multiplicative weights updates algorithm and a conjecture of haploid genetics. arXiv [Preprint] (2014). <http://arxiv.org/abs/1408.6270> (Accessed 6 December 2019).
29. R. Meir, D. Parkes, On sex, evolution, and the multiplicative weights update algorithm. arXiv [Preprint] (2015). <https://arxiv.org/abs/1502.05056v1> (Accessed 6 March 2021).
30. J. O. Campbell, Universal Darwinism as a process of Bayesian inference. *Front. Syst. Neurosci.* **10**, 49 (2016).
31. D. Czigel, H. Giffar, J. B. Tenenbaum, E. Szathmáry, Bayes and Darwin: How replicator populations implement Bayesian computations. *BioEssays* **44**, 2100255 (2022).
32. C. Adami, *The Evolution of Biological Information: How Evolution Creates Complexity, from Viruses to Brains* (Princeton University Press, 2024).
33. N. Shea, Representation in the genome and in other inheritance systems. *Biol. Philos.* **22**, 313–331 (2007).
34. C. T. Bergstrom, M. Rosvall, The transmission sense of information. *Biol. Philos.* **26**, 159–176 (2011).
35. R. S. McGee, "Natural selection as the process of accumulating adaptive information," Ph.D. thesis, University of Washington, United States-Washington (2021).
36. R. A. Fisher, *The Genetical Theory of Natural Selection* (Clarendon Press, Oxford, England, 1930).
37. S. A. Frank, Natural selection. V. How to read the fundamental equations of evolutionary change in terms of information theory. *J. Evol. Biol.* **25**, 2377–2396 (2012).
38. C. Adami, The use of information theory in evolutionary biology. *Ann. N. Y. Acad. Sci.* **1256**, 49–65 (2012).
39. J. Maynard Smith, *Evolution and the Theory of Games* (Cambridge University Press, 1982).
40. I. Eshel, M. W. Feldman, Initial increase of new mutants and some continuity properties of ESS in two-locus systems. *Am. Nat.* **124**, 631–640 (1984).
41. P. Hammerstein, Darwinian adaptation, population genetics and the streetcar theory of evolution. *J. Math. Biol.* **34**, 511–532 (1996).
42. A. Kaznatcheev, Evolution is exponentially more powerful with frequency-dependent selection. bioRxiv [Preprint] (2020). <https://doi.org/10.1101/2020.05.03.075069> (Accessed 6 March 2021).
43. J. F. Crow, Some possibilities for measuring selection intensities in man. *Hum. Biol.* **30**, 1–13 (1958).
44. J. F. Crow, "Genetic loads and the cost of natural selection" in *Mathematical Topics in Population Genetics*, K. I. Kojima, Ed. (Springer, Berlin, 1970), pp. 128–177.
45. M. Kimura, Evolutionary rate at the molecular level. *Nature* **217**, 624–626 (1968).
46. J. Bertram, Z. Shafiei, Strong amplification of quantitative genetic variation under a balance between mutation and fluctuating stabilizing selection. *Genetics* **233**, iyag063 (2026).
47. Y. Freund, R. E. Schapire, Adaptive game playing using Multiplicative Weights. *Games Econom. Behav.* **29**, 79–103 (1999).
48. N. Cesa-Bianchi, G. Lugosi, *Prediction, Learning, and Games* (Cambridge University Press, 2006).
49. S. Arora, E. Hazan, S. Kale, The Multiplicative Weights Update Method: A meta-algorithm and applications. *Theory Comput.* **8**, 121–164 (2012).
50. A. Kaznatcheev, Two conceptions of evolutionary games: Reductive vs effective. bioRxiv [Preprint] (2017). <https://doi.org/10.1101/231993> (Accessed 6 March 2021).
51. C. Bohm, V. R. Ragusa, C. Ofria, R. E. Lenski, C. Adami, Reduced selection during sweeps lead to adaptive momentum on rugged landscapes. bioRxiv [Preprint] (2024). <https://doi.org/10.1101/2024.04.08.588357> (Accessed 15 July 2026).
52. A. Kaznatcheev, Computational complexity as an ultimate constraint on evolution. *Genetics* **212**, 245–265 (2019).
53. H. A. Lindsey, J. Gallie, S. Taylor, B. Kerr, Evolutionary rescue from extinction is contingent on a lower rate of environmental change. *Nature* **494**, 463–467 (2013).
54. R. S. McGee, O. Kosterlitz, A. Kaznatcheev, B. Kerr, C. T. Bergstrom, github.com/ryansmcgee/cost-information-selection: v2023.11.21 (2023). 10.5281/zenodo.10183554. Accessed 21 November 2023.