Optimal evolutionary control for artificial selection on molecular phenotypes

Armita Nourmohammad*
Department of Physics, University of Washington, 3910 15th Ave Northeast, Seattle, WA 98195
Max Planck Institute for Dynamics and Self-organization, am Faßberg 17, 37077 Göttingen, Germany

Ceyhun Eksin
Department of Industrial and Systems Engineering,
Texas A&M University, College Station, TX 77845

Controlling an evolving population is an important task in modern molecular genetics, including in directed evolution to improve the activity of molecules and enzymes, breeding experiments in animals and in plants, and in devising public health strategies to suppress evolving pathogens. An optimal intervention to direct evolution should be designed by considering its impact over an entire stochastic evolutionary trajectory that follows. As a result, a seemingly suboptimal intervention at a given time can be globally optimal as it can open opportunities for more desirable actions in the future. Here, we propose a feedback control formalism to devise globally optimal artificial selection protocol to direct the evolution of molecular phenotypes. We show that artificial selection should be designed to counter evolutionary tradeoffs among multi-variate phenotypes to avoid undesirable outcomes in one phenotype by imposing selection on another. Control by artificial selection is challenged by our ability to predict molecular evolution. We develop an information theoretical framework and show that molecular time-scales for evolution under natural selection can inform how to monitor a population to acquire sufficient predictive information for an effective intervention with artificial selection. Our formalism opens a new avenue for devising optimal artificial selection for directed evolution of molecular functions.

I. INTRODUCTION

The concept of feedback control in molecular evolution was first advocated by *A. Wallace* as a way of describing natural selection [1]. Wallace hypothesized that similar to the centrifugal governor of the steam engine, the action of natural selection is like a *controller* that balances organismic traits, such that weak feet are often accompanied with powerful wings [1]. Such evolutionary tradeoffs are ubiquitous in natural fitness landscapes. For example, experiments on a protein transport system has shown that the fitness landscape for the underlying biochemical network is tuned to exploit optimal control with feedback throughout evolution [2]. However, it remains to be determined whether these structures are solely reflective of biochemical constraints or have emerged as incidences of fitness landscapes that could accommodate for long-term evolutionary survival.

Evolution as a feedback control is also reflected in the inheritance strategies and phenotypic response of populations to time-varying environments. A prominent example of such adaptation is observed in bacteria where cells use phenotypic switches to produce slowly replicating bacteria with tolerance and persistence against antibiotics. Populations use this Lamarckian-type phenotypic response [3] to hedge their bets against fluctuating environments [4, 5]— an optimal response that can be viewed as an evolutionary feedback control [6].

Another approach to evolutionary control is through external interventions with artificial selection to direct populations to acquire a desired trait. Fig. 1 demonstrates artificial selection with a feedback control to breed "pink cows", which are otherwise not favored by natural selection. Such selective breeding has long been used to domesticate animals or to improve agricultural yield in crops and became the basis for Mendelian genetics [7].

Another important avenue for artificial selection is to characterize intervention protocols against rapidly evolving pathogens, for example to counter emergence of drug resistance in bacteria, escape of viral populations from immune challenge, or progression of evolving cancer tumors [8]. Artificial selection also plays a significant role in improving molecular functions. Importantly, directed evolution in the lab is currently being employed to improve the activity and selectivity of molecules and enzymes [9–11], often desirable in industry or for pharmaceutical purposes.

Designing any artificial selection protocol is limited by our ability to predict the outcome of evolution, which is often challenging due to a multitude of stochastic forces at play, such as mutations, reproductive stochasticity (genetic drift) and environmental fluctuations that shape the molecular composition of a population [12, 13]. In contrast to strongly divergent evolution at the genetic level, there is growing experimental evidence for convergent predictable evolution at the phenotypic level [14–16], including for complex molecular phenotypes like RNA polymerase function [17]. We will exploit this evolutionary predictability and focus on designing artificial selec-

*correspondence should be addressed to Armita Nourmohammad: armita@uw.edu

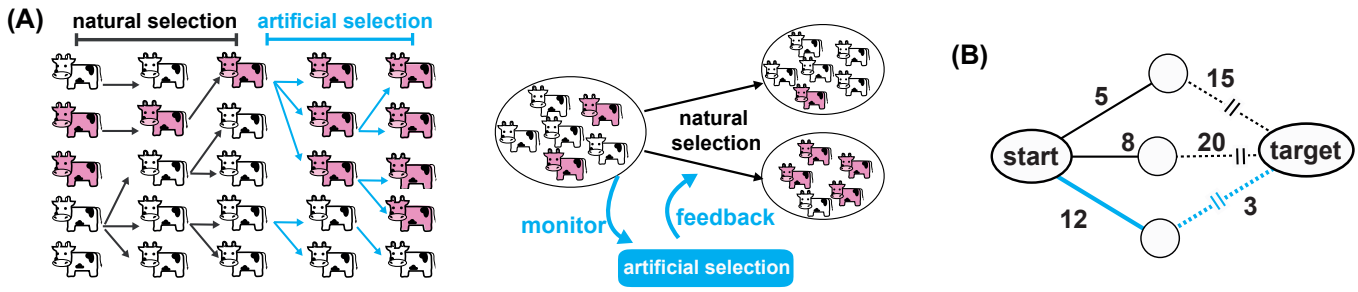

FIG. 1: **Artificial selection as an optimal stochastic adaptive control strategy.** **(A)** Artificial selection is an external intervention to select for a desired trait (i.e., pinkness of cows) in a population, which is otherwise not favored by natural selection. Artificial selection should be viewed as an optimal feedback control, whereby monitoring a stochastically evolving population informs the intervention protocol to optimally bias breeding and reproduction over generations. **(B)** The graph shows different paths with indicated costs for a system to evolve from a start to a target state. Bellman's principle of optimality states that at each step an optimal decision is made, assuming that the following steps are also determined optimally. Although the first step (full line) of the blue path is more costly compared to the others (dotted lines), its cumulative cost is minimum, and hence, it should be chosen as the optimal path. This decision can be made best recursively, known algorithmically as dynamic programming.

tion for molecular phenotypes, which are key links between genotypic information, organismic functions, and evolutionary fitness [12].

Fitness and large-scale organismic traits are often encoded by a number of co-varying molecular phenotypes, linked through genetic interactions; pigmentation patterns on the wings or body of fruit flies are among such multi-dimensional traits, impacted by the expression level of many interacting genes.

A central issue in designing artificial selection for multi-variate phenotypes is to avoid the undesirable (side)effects of selection, which can arise due to evolutionary tradeoffs, e.g. between function and stability of a protein. Evolutionary interventions on multi-variate phenotypes should be designed by assuming their impact over an entire evolutionary trajectory that follows. As a result, a locally optimal action at a given time point may be sub-optimal once considering all the actions that are required to follow in order to direct the correlated evolution of the phenotypes towards their targets; see Fig. 1B.

Finding a globally optimal protocol to direct a stochastic evolution is a topic of control theory, known for its impact in engineering, economics and other fields [18]. Here, we introduce a formalism based on optimal control to devise a population-level artificial selection strategy and drive the stochastic evolution of multi-variate molecular phenotypes towards a desired target. We will show how uncertainty and lack of evolutionary predictability can limit the efficacy of such artificial selection. Finally, we will discuss how to best monitor a population and acquire a sufficient predictive information in order to optimally intervene with its evolution.

II. RESULTS

A. Model of multi-variate phenotypic evolution

Molecular phenotypes are often polymorphic due to genetic variation in their encoding sequence within a population. Here, we primarily focus on phenotypes that are encoded by a relatively large number of genetic loci and hence, are approximately normally distributed within a population— this Gaussian approximation however, can be relaxed as discussed in ref. [19]. In the case of normally distributed k-dimensional phenotypes, we characterize the population's phenotype statistics by the average $\mathbf{x} = [x_1, x_2, \ldots, x_k]^\top$ and a symmetric covariance matrix K, where the diagonal elements $K_{ii}(\mathbf{x})$ indicate the variance of the ith phenotype and the off-diagonal entries $K_{ij}(\mathbf{x})$ indicate the covariance between different phenotypes.

The primary evolutionary forces that shape the composition of phenotypes within a population are selection, mutations and genetic drift at the constitutive genetic sites. Molecular phenotypes are often encoded in confined genomic regions of about a few 100 bps, and hence, are not strongly impacted by recombination, even in sexually reproducing populations. The impact of the evolutionary forces on phenotypes can be directly projected from the evolutionary dynamics in the high-dimensional space of the encoding genotypes [20, 21]. The effect of selection on the mean phenotype is proportional to the covariance between fitness and phenotype within a population [22]. For Gaussian distributed phenotypes, the change in mean phenotype $d\mathbf{x}$ over a short time interval dt simplifies to a stochastic process [23],

$$d\mathbf{x} = (K \cdot \nabla F + \nabla M)dt + \frac{1}{N}\Sigma \cdot d\mathbf{W} \qquad (1)$$

where, F and M are fitness and mutation potentials, re-

spectively. The gradient functions (denoted by ∇F and ∇M) determine the forces acting on the phenotypes by selection and mutation, respectively. $d\mathbf{W}$ is a differential that the reflects the stochastic effect of genetic drift by a multi-dimensional Wiener noise process [24]. The amplitude of the noise is proportional to Σ, which is the square root of the covariance matrix (i.e., $\Sigma^\top \Sigma \equiv K$), scaled by the effective population size N that adjusts the overall strength of the noise. In other words, the fluctuations of the mean phenotype across realizations of an evolutionary process is proportional to the intra-population variance K and inversely scales with the effective population size (i.e., the sample size) N.

Most of our analyses are applicable to general fitness and mutation landscapes. However, we characterize in detail the features of artificial selection to direct evolution on quadratic fitness and mutation landscapes, where phenotypes evolve by natural selection towards an evolutionary optimum [25]. In this case, the impacts of selection and mutation follow linear functions in the high-dimensional phenotypic space, $\nabla F = -2C_0 \cdot \mathbf{x}$, $\nabla M = -2L \cdot \mathbf{x}$, where $\mathbf{x}$ denotes the shifted phenotype vector centered around its stationary state and C_0 and L are selection and mutation strengths, respectively. We can formulate the evolution of mean phenotypes by,

$$d\mathbf{x} = -2KC\,\mathbf{x}\,dt + \Sigma\,d\mathbf{W} \tag{2}$$

where $C \equiv N(C_0 + K^{-1}L)$ is the adaptive pressure, scaled by the population size, which quantifies the potential of a phenotype to accumulate mutations under selection. In the rest of the manuscript, we will use F as a short hand for the adaptive landscape under natural selection, whose gradient characterizes the adaptive pressure, $\nabla F = -2C\mathbf{x}$ in eq. 2. We have also rescaled time with the effective population size (i.e., $t \to Nt$), which is the coalescence time in neutrality [26].

Similar to the mean, the covariance matrix K is a time-dependent variable, impacted by evolution. However, fluctuations of covariance are much faster compared to the mean phenotype, and therefore, covariance can be approximated by its stationary ensemble-averaged estimate [19, 20]. Moreover, even in the presence of moderately strong selection pressure, the phenotypic covariance depends only weakly on the strength of selection and is primarily determined by the supply of mutations in a population [20, 21]. Therefore, we also assume that the phenotypic covariance matrix remains approximately constant over time, throughout evolution. With these approximations, evolution of the mean phenotype can be described as a stochastic process with a constant adaptive pressure that approaches its stationary state over a characteristic equilibration time $\sim (2KC)^{-1}$.

The stochastic evolution of the mean phenotype in eq. 2 defines an ensemble of evolutionary trajectories. We can characterize the statistics of these evolutionary paths by the dynamics of the underlying conditional probability density $P(\mathbf{x}', t'|\mathbf{x}, t)$ for a population to have a mean phenotype $\mathbf{x}'$ at time t', given its state $\mathbf{x}$ at an earlier time $t < t'$. The dynamics of this probability density follows a high-dimensional Fokker-Planck equation [20],

$$\frac{\partial}{\partial t} P(\mathbf{x}', t'|\mathbf{x}, t) = \left[\frac{1}{2N}\mathbf{Tr} K \nabla_{\mathbf{xx}} - \nabla(K \cdot \nabla F)\right] P(\mathbf{x}', t'|\mathbf{x}, t) \tag{3}$$

where we introduced the compact notation, $\mathbf{Tr} K \nabla_{\mathbf{xx}} \equiv \sum_{ij} K_{ij} \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}$. As a result, the conditional distribution of phenotypes follows an Ornstein-Uhlenbeck process, described by a time-dependent multi-variate Gaussian distribution; see Appendix A.

B. Artificial selection to optimally direct evolution

Natural selection in eqs. 2,3 drives populations towards an optimum, which in general is a function of the organism's environmental and physiological constraints. Artificial selection aims to relax or tighten some of the natural constraints to drive evolution towards an alternative desired state $\mathbf{x}^*$. In general, we can formulate evolution subject to artificial selection as,

$$d\mathbf{x} = \underbrace{\Big(K \cdot \nabla F + \mathbf{u}(\mathbf{x}, t)\Big)}_{A(\mathbf{x},t)} dt + \Sigma d\mathbf{W} \tag{4}$$

where $\mathbf{u}(\mathbf{x}, t)$ is a time- and phenotype-dependent vector, which denotes the impact of artificial selection and $A(\mathbf{x}, t)$ is the total force incurred by natural and artificial selection on the phenotypes. Our goal is to find an optimal protocol for artificial selection $\mathbf{u}(\mathbf{x}, t)$ in order to reach the target $\mathbf{x}^*$ by a desired time t_f, while minimizing the cost function,

$$\Omega(\mathbf{x}, \mathbf{u}, t) = V(\mathbf{x}, t) + \frac{1}{2}\mathbf{u}^\top B \mathbf{u} \tag{5}$$

over an entire evolutionary trajectory. Here, $V(\mathbf{x}, t) \equiv V(|\mathbf{x}_t - \mathbf{x}^*|)$ is the cost for deviation of the phenotype state $\mathbf{x}_t$ at time t from the desired target $\mathbf{x}^*$, and B is a matrix that characterizes the cost for imposing artificial selection $\mathbf{u} \equiv \mathbf{u}(\mathbf{x}, t)$ and intervening with natural evolution of each phenotype.

Finding an optimal strategy $\mathbf{u}^*(\mathbf{x}, t)$ to steer a stochastic process is a topic of control theory, known for its impact in engineering, economics and other fields [18]. To solve the optimal control problem (i.e., to characterize an optimal artificial selection strategy), we define the *cost-to-go* function,

$$J(\mathbf{x}, t) = \min_{\mathbf{u}} \left\langle Q(\mathbf{x}, t_f) + \int_t^{t_f} ds \left(V(\mathbf{x}_s) + \frac{1}{2}\mathbf{u}_s^\top B \mathbf{u}_s\right)\right\rangle_{\text{evol.}} \tag{6}$$

where the angular brackets $\langle\cdot\rangle$ indicate expectation over stochastic evolutionary histories from time t until the target time t_f. Here, $Q(\mathbf{x}, t_f) \equiv Q(|\mathbf{x}_{t_f} - \mathbf{x}^*|)$ characterizes the cost of deviation from the target at the end of the

evolutionary process t_f, which could be chosen to be different from the path cost $V(\mathbf{x})$.

An optimal artificial selection protocol should be designed by considering its impact over an entire stochastic evolutionary trajectory that follows. As a result, a seemingly suboptimal intervention at a given time can be globally optimal as it can open opportunities for more desirable actions in the future; see schematic Fig. 1B. To characterize a globally optimal artificial selection protocol at each time point $\mathbf{u}^*(\mathbf{x}, t)$, we will assume that the following selection strategies will be also determined optimally. This criteria is known as Bellman's "principle of optimality" [27], and would allow us to express the optimal control problem in a recursive form (Appendix A), known as dynamic programming in computer science [27]. As a result, we can formulate a dynamical equation for the cost-to-go function, known as Hamilton-Jacobi-Bellman (HJB) equation [28],

$$-\frac{\partial J(\mathbf{x},t)}{\partial t} = \min_{\mathbf{u}} \left[\Omega(\mathbf{x}_t, \mathbf{u}_t) + A(\mathbf{x}_t) \cdot \nabla J + \frac{1}{2} \mathbf{Tr} K \nabla_{\mathbf{xx}} J \right] \tag{7}$$

with the boundary condition $J(\mathbf{x}, t_f) = Q(\mathbf{x}, t_f)$ at the end of the process; see Appendix A. Importantly, the HJB equation (7) indicates that the cost-to-go $J(\mathbf{x}, t)$ is a potential function based on which the optimal artificial selection can be evaluated,

$$\mathbf{u}^*(\mathbf{x}, t) = -B^{-1} \cdot \nabla J(\mathbf{x}, t). \tag{8}$$

In other words, the cost-to-go function characterizes a time- and phenotype-dependent artificial fitness landscape that determines the strength of artificial selection $\mathbf{u}^*(\mathbf{x}, t)$.

C. Artificial selection by path integral control

The solution to the HJB equation (7) for the cost-to-go function $J(\mathbf{x}, t)$ and the artificial selection $\mathbf{u}^*(\mathbf{x}, t)$ can be complex time- and state-dependent functions, described by non-linear evolutionary operators (Appendix A). Here, we consider a class of control problems, known as "path integral control" [29–31], where the cost matrix B for artificially intervening with evolution is inversely proportional to the phenotypic covariance K, i.e., $B = \lambda K^{-1}$, where λ is a constant that determines the overall cost of artificial selection. In other words, we assume that imposing artificial selection on highly conserved phenotypes is more costly than on variable phenotypes. This is intuitive as conserved phenotypes are often essential for viability of an organism and it is best to design a control cost function that limits the access to such phenotypes through artificial selection.

Importantly, the path-integral control assumption results in a significant mathematical simplification for the dynamics of the of the cost-to-go function $J(\mathbf{x}, t)$ and makes the inference of optimal artificial selection more tractable; see Appendices A, B. We can characterize the evolution of the conditional distribution $P_u(\mathbf{x}', t'|\mathbf{x}, t)$ for a population under optimal artificial selection $\mathbf{u}^*(\mathbf{x}, t)$ to be in the phenotypic state $\mathbf{x}'$ at time t', given its state $\mathbf{x}$ at time t by,

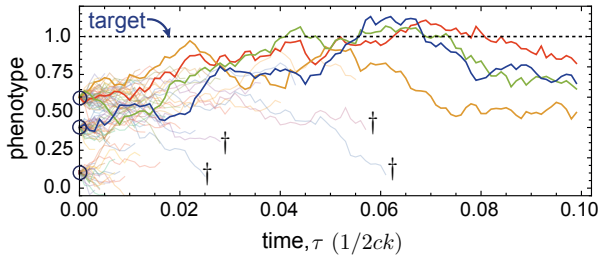

FIG. 2: **Artificial selection with stochastic optimal annihilation.** Phenotypic trajectories starting from three distinct initial states (open circles) are shown for evolution under natural selection in a 1D quadratic landscape of eq. 2. The trajectories are annihilated († and low opacity lines) with a rate proportional to the cost of their deviation from target (dotted line) $V(|x - x^*|)/\lambda$ (eq. 9). At each time point, the survived trajectories characterize the ensemble of evolutionary states for evolution under optimal artificial selection to reach the target at $x^* = 1$. Time is measured in units of the characteristic time for natural evolution $(1/2kc)$. Parameters: $k = 0.4$; $c = 1$; $\lambda = 0.01$; $g = 2$.

$$\begin{aligned} &\frac{\partial}{\partial t} P_u(\mathbf{x}', t'|\mathbf{x}, t) \\ &= \left[\frac{1}{2N} \mathbf{Tr} K \nabla_{\mathbf{xx}} - \nabla(K \nabla F) - \frac{1}{\lambda} V(\mathbf{x}, t) \right] P_u(\mathbf{x}', t'|\mathbf{x}, t) \end{aligned} \tag{9}$$

with the initial condition $P_u(\mathbf{x}', t|\mathbf{x}, t) = \delta(\mathbf{x} - \mathbf{x}')$ (see Appendix A). The cost-to-go function $J(\mathbf{x}, t)$, and consequently the optimal control $\mathbf{u}^*$ (eq. 8), can be directly evaluated as a marginalized form of the conditional probability density in eq. 9 (Appendix A),

$$J(\mathbf{x}, t) = -\lambda \log \int d\mathbf{x}' P_u(\mathbf{x}', t_f|\mathbf{x}, t) \exp[-Q(\mathbf{x}', t_f)/\lambda] \tag{10}$$

where $Q(\mathbf{x}', t_f)$ is the end point cost (eq. 6). Evolution under artificial selection (eq. 9) resembles the natural evolutionary dynamics (eq. 3 with $\mathbf{u} = 0$) with an extra annihilation term $V(\mathbf{x}, t)/\lambda$ [30]. Therefore, artificial selection acts as an *importance sampling* protocol over each selection cycle (e.g. each generation) that removes (annihilates) individuals from the population with a rate proportional to their distance from the evolutionary target $\sim V(|\mathbf{x}_t - \mathbf{x}^*|)/\lambda$; see Fig. 2. Specifically, at each time point, this protocol generates a phenotypic distribution consistent with the evolutionary process under optimal artificial selection in eq. 4, without an explicit knowledge of the selection protocol $\mathbf{u}^*(\mathbf{x}, t)$; see Appendix A.

Moreover, in the path integral control regime (i.e., $B = \lambda K^{-1}$), the cost-to-go function scaled by the overall control factor, $J(\mathbf{x}, t)/\lambda$, determines a time- and phenotype-dependent fitness landscape associated with artificial selection $F_{\text{art.}}(\mathbf{x}, t)$; see eq. 8). Throughout an artificial selection process, populations experience an effective landscape $\hat{F}(\mathbf{x}, t) = F(\mathbf{x}) + F_{\text{art.}}(\mathbf{x}, t)$, as a combination of the natural fitness landscape $F(\mathbf{x})$ and the artificial fitness landscape, $F_{\text{art.}}(\mathbf{x}, t)$. The overall cost of control λ determines the impact of artificial selection on evolution relative to natural selection, and when the control cost is small (i.e., $\lambda \ll 1$), artificial selection can dominate the course of evolution.

D. Artificial selection for phenotypes under stabilizing selection

Here, we focus on the specific case of natural evolution in a high dimensional quadratic fitness landscape (eq. 2). In addition, we assume a quadratic form for the cost function throughout the evolutionary process, $V(\mathbf{x}, t) = \frac{1}{2}(\mathbf{x}_t - \mathbf{x}^*)^\top G(\mathbf{x}_t - \mathbf{x}^*)$ and also at the end point $Q(\mathbf{x}, t_f) = \frac{1}{2}(\mathbf{x}_{t_f} - \mathbf{x}^*)^\top \tilde{G}(\mathbf{x}_{t_f} - \mathbf{x}^*)$. Characterizing an artificial selection protocol under such quadratic constraints falls within the class of standard stochastic control problems, known as linear-quadratic-Gaussian (LQG) control [18]. However, we will present our analyses based on the path integral control approach in eq. 9, which is generalizable beyond LQG and can be applied to arbitrary cost functions and fitness landscapes (see Appendix B).

Let us imagine that our criteria is to drive evolution towards the optimum $\mathbf{x}^*$ by time t_f, which implies that the path cost is zero $G = 0$ but the end-point cost is non-zero $\tilde{G} > 0$; see the above section and Appendix B for the general scenario including the case with $G > 0$. In this case, we infer that the strength of the optimal artificial selection should strongly increase as we approach the end point,

$$\mathbf{u}^*(\mathbf{x}, t) = -\frac{K}{\lambda} e^{-2CK\tau} \tilde{G} \left[I - \frac{K_\tau}{\lambda} \left(I + \frac{K_\tau}{\lambda} \tilde{G} \right)^{-1} \tilde{G} \right] (e^{-2KC\tau}\mathbf{x} - \mathbf{x}^*) \tag{11}$$

where $\tau = t_f - t$ is the remaining time to the end point and $K_\tau = \langle \mathbf{x}(t), \mathbf{x}(t_f) \rangle$ is the time-dependent covariance matrix for the conditional probability density of natural evolutionary process; see Appendix B for the case of $G > 0$ and Appendix C for detailed derivation. At the end point, the optimal artificial selection keeps the population close to the target with a strength,

$$\mathbf{u}^*(\tau \to 0) = -\frac{1}{\lambda} K \tilde{G} (\mathbf{x} - \mathbf{x}^*) \tag{12}$$

which resembles the breeder's equation [32] for artificial selection with a heritability factor, $h^2 = K\tilde{G}/\lambda$.

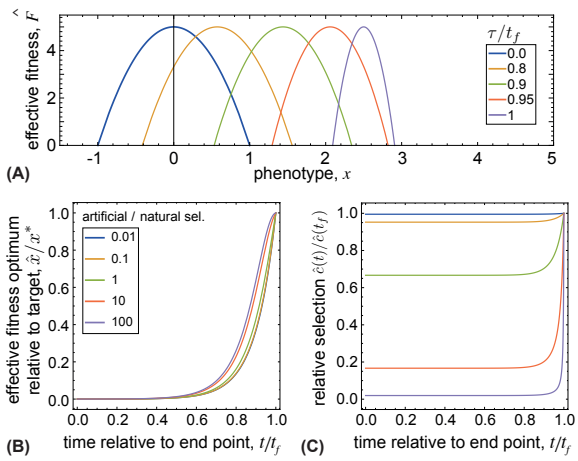

FIG. 3: **Optimal artificial selection for a 1D phenotype.** The impact of artificial selection intensifies as time approaches the end point of the process. **(A)** The interplay between artificial and natural selection defines an effective time-dependent (colors) fitness landscape $\hat{F}$ with an optimum $\hat{x}(t)$ that approaches the phenotypic target for artificial selection ($x^* = 3$) and an effective selection pressure $\hat{c}$ that intensifies as time approaches end point $t/t_f \to 1$. Other parameters: $\lambda = 0.1$; $c = 2$; $g = 2$. **(B)** and **(C)** show the effective fitness peak relative to the target $\hat{x}/x^*$ and the relative selection pressure of the effective fitness landscape $\hat{c}(t)/\hat{c}(t_f)$ as a function of time, for a range of relative artificial to natural selection pressures $g/\lambda c$ (colors).

When the goal is to drive the population towards a target by an end point t_f, the effective fitness $\hat{F}(\mathbf{x}, t)$ remains close to the natural landscape for an extended period. As time approaches the end point, populations transition from evolving in their natural landscape $F(\mathbf{x})$ to the artificially induced fitness landscape $F_{\text{art.}}(\mathbf{x}, t_f)$; see Figs. 3 and 4 for evolution in one and two dimensions, respectively. Moreover, towards the end point, the fitness peak and the strength of selection approach their final values, determined by the target and the cost functions in eq. 5, in an exponentially rapid manner; Fig. 3B,C.

E. Artificial selection for multi-variate phenotypes

One of the main issues in designing breeding experiments in plants and animals is the undesirable (side)effects of artificial selection on covarying phenotypes, primarily due to evolutionary tradeoffs [33] e.g. between sturdiness and flavor of vegetables like tomatoes [34]. Similarly, tradeoffs among covarying molecular phenotypes (e.g. function vs. stability of a protein) could lead to undesirable outcomes for artificial selection at the molecular level.

To demonstrate the consequences of phenotypic covariation, let us consider a simple example for artificial selection on two covarying phenotype (x, y). We aim to drive the phenotype x towards the target $x^* > 0$ by artificial

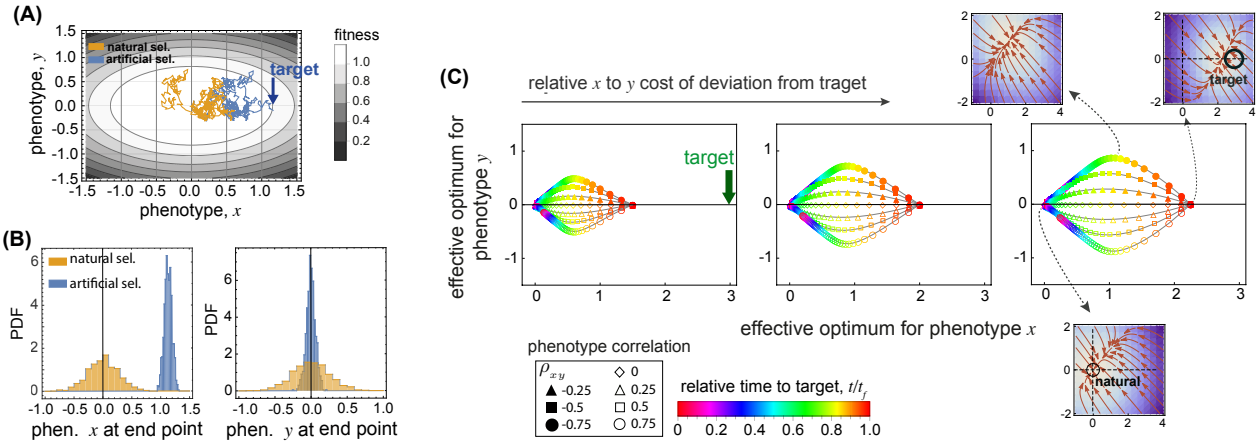

FIG. 4: **Artificial selection on covarying phenotypes. (A)** Trajectories for evolution under natural (orange) and artificial (blue) selection are shown for a 2D phenotype (x, y), in a quadratic landscape. Parameter: $c_x = 2$, $c_y = 4$, $c_{xy} = 0$; $x^* = 1.2$, $y^* = 0$; $k_x = 0.02$; $k_y = 0.05$; $k_{xy} = 0$; $g_x = g_y = 2$; $\lambda = 0.01$. **(B)** The distribution of phenotypes at the end point of an artificial selection protocol (blue) is compared to the phenotypic distribution under natural selection (orange). Evolutionary parameters are the same as in (A). **(C)** The dynamics of the effective fitness peak is shown over time (colors) for 2D covarying phenotypes with correlations ρ_{xy} indicated by the shape of the markers. From left to right, panels show increasing end-point cost of deviation from the target along the x-phenotype, $g_x = 1, 2, 3$, with $g_y = 2$. Heatmaps show the effective fitness landscapes associated with a specific fitness peak (indicated by the dotted arrows) for anti-correlated phenotypes at three different time points. The direction and length of the red arrows in each heatmap indicate the direction and the strength of selection pressure towards the effective optimum. Parameters: $x^* = 3$, $y^* = 0$; $c_x = c_y = 5$, $c_{xy} = 0$; $k_x = k_y = 0.02$; $\lambda = 0.1$.

selection while keeping the phenotype y at its stationary state value $y^* = 0$. An optimal artificial selection protocol (eq. 11) defines an effective two-dimensional quadratic fitness landscape that biases the evolutionary trajectories towards the target state (Fig. 4A). As a result, the phenotype distributions at the end of the process become significantly distinct from the expectation under natural selection, and remain strongly restricted around their targets values; Fig. 4B.

The peak of this fitness landscape (i.e., the effective optimum) changes from the natural state $(0, 0)$ to the target state (x^*, y^*) by the end of the selection process; Fig. 4C and Fig. S1. The fitness peak moves monotonically along the x-phenotype from the natural optimum 0 towards the target x^*, irrespective of the correlation $\rho_{x,y}$ between the two phenotypes; Fig. 4C. However, the dynamics of the fitness peak along the y-phenotype is generally non-monotonic and strongly dependent on the phenotypic correlation $\rho_{x,y}$. An increase in x drives the positively (negatively) correlated y towards higher (lower) values. Therefore, in the beginning of the process, the optimal artificial selection protocol sets the fitness peak for the y-phenotype at an opposite direction to counter-balance the effect of evolutionary forces due to phenotypic covariation. As the end-point approaches, artificial selection becomes significantly strong with an effective fitness optima set at the target for each phenotype x^* and y^* (eq. 12). Therefore, the optimum y-value should return to its target state ($y^* = 0$), resulting a non-monotonic behavior in the dynamics of the fitness peak along the y-phenotype; see Fig. 4C. Moreover, the strength of selection also transitions over time and becomes stronger towards the target phenotypes at the end point (heatmaps in Fig. 4C and Fig. S2).

The optimal artificial selection protocol in Fig. 4 is strongly sensitive to the structure of the phenotypic covariance matrix K (eq. 11), and hence, any protocol should be carefully designed to achieve the intended phenotypic desirability.

F. Artificial selection with intermittent monitoring

Imposing artificial selection based on continuous feedback from the state of the population (Fig. 1) requires complete monitoring and the possibility of continuous evolutionary intervention— a criteria that is often not met in real conditions. In general, discrete selection protocols based on a limited information can be inefficient for devising evolutionary feedback control [8, 35]. Here, we characterize the limits of discrete optimal interventions based on the evolutionary response of the population to artificial selection. We consider a simple scenario where in a stationary state we aim to keep a population at the target phenotype $\mathbf{x}^*$, using discrete monitoring and interventions at time points $(i = 1, \ldots, M)$ with a time separation $\tau \equiv t_{i+1} - t_i$. We define a stationary cost-to-go function,

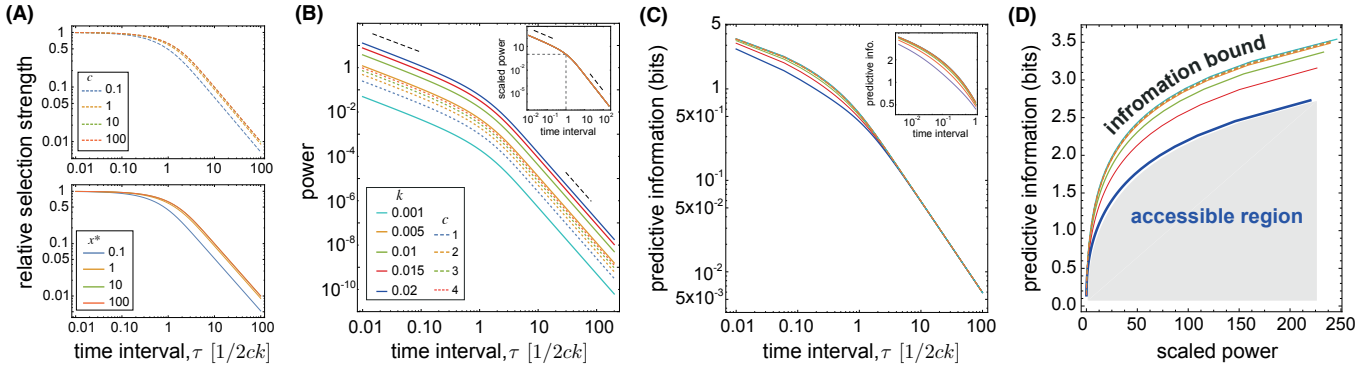

FIG. 5: **Artificial selection with limited information.** **(A)** Relative strength of artificial selection α_τ/α_0 (eq. 14) is shown as a function of the time interval for monitoring and intervention τ, measured in units of the characteristic time for evolution under natural selection $(1/2kc)$. The selection ratio is shown for various strengths of natural selection c (top; with $x^* = 1$) and for various targets of artificial selection x^* (bottom; with $c = 1$). **(B)** Power (eq. 16) is shown as a function of the time interval τ for a range of parameters for the phenotypic diversity k (full line) and the strength of natural selection c (dotted line). The insert shows a collapse plot for power scaled with the expectation at the characteristic time for natural selection $\text{Power}(\tau)/\text{Power}(\tau = 1)$. **(C)** Predictive mutual information $\mathcal{I}(\tau)$ (eq. 17) is shown to decay with the time interval τ for a wide range of parameters (k, c). Insert zooms into a narrower range for the time interval $\tau < 1$. **(D)** Predictive information (eq. 17) is shown as a function of the scaled power for optimized artificial selection (eq. 16). Each curve sets an information bound for artificial selection for a given set of evolutionary parameters (k, c). A non-optimal selection intervention should lie below the information curve, shown by the gray shaded area as the accessible region associated with the dark blue curve. Color code in (C,D) is similar to (B). Other parameters: $\lambda = 0.6$; $x^* = 3$; $g = 2$.

$$J(\mathbf{x}, t_m; \tau) = \min_{\mathbf{u}} \lim_{M\to\infty} \frac{1}{(M-m)\tau} \left\langle \sum_{i=m}^{M} \mathbf{u}_i^\top B \mathbf{u}_i + \int_{t_i}^{t_M} V(\mathbf{x}_t) dt \right\rangle_{\text{evol.}} \tag{13}$$

where the devision by the total time $M\tau$ assures that the cost-to-go remains finite. To further simplify, we only consider one dimensional phenotype x with intra-population variance k, the cost of deviation $V(x) = g(x - x^*)^2/2$ from target x^*, and the cost of intervention $\beta u^2/2$ with artificial selection u. However, our analyses can be easily generalized to multi-variate phenotypes.

In the stationary state and in the regime of small perturbations $(gk/\lambda \ll 1)$, the optimal artificial selection protocol u^* should be a variant of the case with full information with a strength of selection α_τ dependent on the time window τ, $u^*_\tau = -k\alpha_\tau(x - x^*)$; see Appendix D. We can characterize the optimal strength of artificial selection α_τ by minimizing the cost-to-go function in eq. 13,

$$\alpha_\tau = \alpha_0 \left[\frac{1/(4c)(1 - e^{-2\tau}) + 2(x^*)^2(1 - e^{-\tau})}{\tau/(2c)(1 + 4c(x^*)^2)} \right] \tag{14}$$

where $\alpha_0 = g/\lambda$ is the optimal selection strength under continuous monitoring. Here, time τ is measured in units of the characteristic time for evolution under natural selection, i.e., $(2kc)^{-1}$.

The partially informed artificial selection α_τ depends on most of the evolutionary parameters similar to selection with complete monitoring α_0. Importantly, the ratio α_τ/α_0, depends only weakly on the strength of natural selection c (Fig. 5A; top) and the target for artificial selection x^* (Fig. 5A; bottom) and it is insensitive to the phenotypic diversity k and the parameter λ (eq. 14).

However, the optimal artificial selection α_τ strongly depends on the time interval τ and it decays as the time interval τ broadens (Fig. 5A). This decay is linear and relatively slow up to the characteristic time for evolution under natural selection $(2kc)^{-1}$. This is the time scale over which an intermittent artificial selection can still contain the population around the desired target x^*. If interventions are separated further in time (i.e., $\tau \gg 1$), the optimal selection strength decays rapidly as $\sim \tau^{-1}$. Imposing a strong artificial selection in this regime is highly ineffective as populations can easily drift away from the target and towards their natural state within each time interval, and any artificial selection would only contribute to the intervention cost $\sim u^2$ without offering any benefits.

G. Information cost for artificial selection

Artificial selection is an evolutionary drive that shifts the equilibrium state of the population under natural selection to a new state around the target. As a result, the phenotypic histories $\mathbf{x}_{t_0,\ldots t_f}$ over the period of $(t_0, \ldots, t_f)$ are statistically distinct for evolution under natural and artificial selection (Fig. 4A). This deviation can be quantified by the Kullback-Leibler distance $D_{KL}(\mathcal{P}_u(\mathbf{x})||\mathcal{P}(\mathbf{x}))$ between the distribution of histories under artificial selection $\mathcal{P}_u(\mathbf{x}) \equiv \mathcal{P}_u(\mathbf{x}_{t_0,\ldots t_f})$ and un-

der natural selection $\mathcal{P}(\mathbf{x})$. In the stationary state, the Kullback-Leibler distance quantifies the impact of artificial selection on evolution and can be interpreted as the amount of work $W_{t_0}^{t_f}(\mathbf{u})$ done by external forces [36] (i.e., the artificial selection) to shift the population from its natural equilibrium to the artificial equilibrium,

$$\begin{aligned} W_{t_0}^{t_f}(\mathbf{u}) &= D_{KL}(\mathcal{P}_u(\mathbf{x})||\mathcal{P}(\mathbf{x})) \\ &= \int d\mathbf{x}_{t_0}^{t_f}\, \mathcal{P}_u(\mathbf{x}) \log\left[\frac{\mathcal{P}_u(\mathbf{x})}{\mathcal{P}(\mathbf{x})}\right] \end{aligned} \tag{15}$$

The cumulative work is related to the cost of artificial selection, and for the case of path integral control, it is equivalent to the cumulative cost of control $W_{t_0}^{t_f}(\mathbf{u}) = \langle \frac{1}{2}\int \mathbf{u}^\top K^{-1}\mathbf{u}\, dt\rangle = \frac{1}{2\lambda}\langle \int \mathbf{u}^\top B\mathbf{u}\, dt\rangle$, where the angular brackets $\langle\cdot\rangle$ denote expectation over the ensemble of evolutionary histories under artificial selection; see refs. [37, 38] and Appendix E. The power (i.e., work per unit time), associated with intermittent artificial selection, can be expressed as the amount of work per time interval τ,

$$\text{Power}(\tau) = \lim_{M\to\infty} \frac{1}{M\tau}\sum_{i=1}^{M} W(t_i) = \frac{1}{2\tau}\langle \mathbf{u}_\tau^\top K^{-1}\mathbf{u}_\tau\rangle \tag{16}$$

The expected work, and hence the power, depend on the time interval τ through various factors. Work depends quadratically on the strength of artificial selection α_τ and on the expected population's deviation from the target $\langle (x_\tau - x^*)^2\rangle$. On the one hand, the optimal strength of artificial selection α_τ decays with increasing the time interval; see Fig. 5A and eq. 14. On the other hand, as the time interval broadens, populations deviate from the target towards their natural state, resulting in an increase in the expected work by artificial selection. Moreover, since interventions are done once per cycle, the power has an overall scaling with the inverse of the cycle interval $\sim \tau^{-1}$. These factors together result in a reduction of the expected power associated with artificial selection as the time interval widens; see Fig. 5B.

Power depends strongly on the parameters of natural evolution including the strength of natural selection (c) and the phenotypic diversity within a population (k); Fig. 5B. This is due to the fact that steering evolution under strong natural selection (i.e., with large k and c) is more difficult and would require a larger power by artificial selection. However, the dependence of power on the evolutionary parameters (k, c) remain approximately separated from its dependence on the time interval τ. Thus, power rescaled by its expectation at the characteristic time $\tau = 1$ shows a universal time-decaying behavior, independent of the evolutionary parameters (Fig. 5B).

H. Predictive information as a limit for efficient artificial selection

Artificial selection can only be effective to the extent that an intervention is predictive of the state of the population in the future. The mutual information between artificial selection and the future state of the population quantifies the amount of predictive information [39] by artificial selection, or alternatively, the memory of the population from the selection intervention. We characterize the predictive information $\mathcal{I}_\tau$ as a time-averaged mutual information $I(x_t, x_0)$ between an intervention (at time $t = 0$) and the state of the population at a later time t, $(0 < t < \tau)$, during each intervention cycle in the stationary state,

$$\begin{aligned} \mathcal{I}_\tau &= \frac{1}{\tau}\int_0^\tau dt\, I(x_t, x_0) \\ &= \frac{1}{\tau}\int dt \int dx_0 dx_t P(x_t, x_0) \log\left[\frac{P(x_t|x_0)}{P(x_t)}\right] \end{aligned} \tag{17}$$

The predictive mutual information monotonically decreases as the time interval τ increases and the population evolves away from the selection target; see Fig. 5C.

Predictive information in eq. 17 quantifies the impact of artificial selection on the future of a population. The information theoretical measure of power in eq. 16 on the other hand, quantifies how the optimal artificial selection protocol distinguishes a population's past evolutionary history from the expectation under natural selection. The efficacy of any intervention (i.e., power) is tightly bounded by the impact it may have on the state of the population in the future (i.e., predictive information); see Fig. 5D. Any non-optimal artificial selection effort should lie below this bound and within the accessible region of the information-power plane (Fig. 5D).

Phenotypic diversity k characterizes the rate at which a population evolves away from the target and towards its natural state during an intervention interval (eq. 2). As a result, the information bound for artificial selection is tighter in more diverse populations, which can rapidly evolve away from the target and towards their natural state during each time interval τ.

As interventions become more frequent, predictive mutual information increases but more slowly than the amount of power necessary to induce an effective artificial selection (Fig. 5D). Importantly, the gain in predictive information becomes much less efficient for time intervals shorter than the characteristic time of natural selection ($\tau \ll 1$).

Trading power with information provides a guideline for scheduling selection interventions. The characteristic time for evolution under natural selection is a powerful gauge for scheduling the interventions. Importantly, setting the time interval within the range of the characteristic evolutionary time $\tau \sim 1$ could provide a sufficient power-to-information ratio for an optimal artificial se-

lection protocol. However, as information becomes less predictive or the inferred selection protocol becomes less optimal, it would be necessary to monitor and intervene more frequently.

III. DISCUSSION

An optimal intervention should be designed by considering its impact over an entire evolutionary trajectory that follows. Here, we infer an artificial selection strategy as an optimal control with feedback to drive multivariate molecular phenotypes towards a desired target. This selection protocol is optimal over the defined time course and may seem sub-optimal on short time-scales as it drives one phenotype away from its target while driving another towards the target to overcome tradeoffs (Fig. 4C). Monitoring and feedback from the state of a population are key for imposing an effective artificial selection strategy. We show that the schedule for monitoring should be informed by the molecular timescales of evolution under natural selection, which set the rate at which a population loses memory of artificial interventions by evolving towards its favorable state under natural selection.

Being able to control evolution could have significant applications in designing novel molecular functions or in suppressing the emergence of undesirable resistant strains of pathogens or cancers. Interventions that select for desired phenotypes have become possible in molecular engineering [40, 41] and in targeted immune-based therapies against evolving pathogens [42]. However, the efficacy of these actions are limited by our ability to monitor and predict the evolutionary dynamics in response to interventions.

Evolution is shaped by a multitude of stochastic effects, including the stochasticity in the rise of novel beneficial mutations and fluctuations in the environment, which at best, raise skepticism about predicting evolution [12, 13]. However, evolutionary predictability is not an absolute concept and it depends strongly on the time window and the molecular features that we are interested in. For example, despite a rapid evolutionary turnover in the influenza virus, a number of studies have successfully forecasted the dominant circulating strains for a one year period [43, 44]. Similarly, phenotypic convergence across parallel evolving populations has been reported as an evidence for phenotypic predictability, despite a wide genotypic divergence [14–17]. Therefore, to exploit the evolutionary predictability for the purpose of control, it is essential to identify the relevant degrees of freedom (e.g., phenotypes vs. genotypes) that confer predictive information and to characterize the evolutionary time scales over which our observations from a population can inform our interventions to drive the future evolution.

We focus on modeling control of molecular phenotypes. Phenotypic diversity within a population provides standing variation that selection can act upon. To allow for a continuous impact of artificial selection over many generations, we have limited our analyses to a regime of moderate time- and phenotype-dependent artificial selection to sustain the phenotypic variability in a population. However, it would be interesting to allow for stronger artificial selection to significantly impact the standing variation and the structure of the phenotypic covariance within a population over time. Indeed, allowing a population to rapidly collapse as it approaches a desired target is a common strategy in evolutionary optimization algorithms [45]— a strategy that could accelerate the design of new functions with directed molecular evolution.

In this work, we assume a stochastic model for evolution of multi-variate molecular phenotypes, which has been powerful in describing a range biological systems, including the evolution of gene expression levels [46]. Indeed, optimal control protocols are often designed by assuming a specific model for the underlying dynamics. However, in most biological systems, we lack a knowledge of the details and the relevant parameters of the underlying evolutionary process. The ultimate goal is to simultaneously infer an effective evolutionary model based on the accumulating observations and to design an optimal intervention to control the future evolution— an involved optimization problem known as dual adaptive control [47].

Acknowledgements

This work has been supported by the DFG grant (SFB1310) for Predictability in Evolution (AN, CE) and the MPRG funding through the Max Planck Society (AN).

[1] C. R. Darwin and A. R. Wallace, Journal of the Proceedings of the Linnean Society of London **Zoology 3**, 45 (1858).
[2] R. Chakrabarti, H. Rabitz, S. L. Springs, and G. L. McLendon, Phys. Rev. Lett. **100**, 258103 (2008).
[3] J. B. de Lamarck, *Philosophie zoologique; ou, Exposition des considérations relatives à l'histoire naturelle des animaux.* , Paris: Dentu (1809).
[4] N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, and S. Leibler, Science **305**, 1622 (2004).
[5] K. Paarporn, C. Eksin, and J. S. Weitz, Journal of theoretical biology **454**, 376 (2018).
[6] O. Rivoire and S. Leibler, Proc. Natl. Acad. Sci. U.S.A. **111**, E1940 (2014).
[7] G. Mendel, Verhandlungen des naturforschenden Vereines in Brünn **5**, 3 (1862).
[8] A. Fischer, I. Vazquez-Garcia, and V. Mustonen, Proc. Natl. Acad. Sci. U.S.A. **112**, 1007 (2015).
[9] M. Eigen and W. Gardiner, Pure Appl. Chem. **56**, 967 (1984).
[10] K. Chen and F. H. Arnold, Proc. Natl. Acad. Sci. U.S.A. **90**, 5618 (1993).

[11] K. M. Esvelt, J. C. Carlson, and D. R. Liu, Nature **472**, 499 (2011).
[12] A. Nourmohammad, T. Held, and M. Lässig, Curr Opin Genet Dev **23**, 684 (2013).
[13] M. Lässig, V. Mustonen, and A. M. Walczak, Nat Ecol Evol **1**, 77 (2017).
[14] E. Toprak, A. Veres, J.-B. Michel, R. Chait, D. L. Hartl, and R. Kishony, Nat. Genet. **44**, 101 (2012).
[15] S. Kryazhimskiy, D. P. Rice, E. R. Jerison, and M. M. Desai, Science **344**, 1519 (2014).
[16] J. Barroso-Batista, J. Demengeot, and I. Gordo, Nat Commun **6**, 8945 (2015).
[17] O. Tenaillon, A. Rodríguez-Verdugo, R. L. Gaut, P. McDonald, A. F. Bennett, A. D. Long, and B. S. Gaut, Science **335**, 457 (2012).
[18] D. P. Bertsekas, *Dynamic programming and optimal control* (Belmont, Mass. : Athena Scientific, 1995).
[19] A. Nourmohammad, J. Otwinowski, and J. B. Plotkin, PLoS Genet **12**, e1006171 (2016).
[20] A. Nourmohammad, S. Schiffels, and M. Lässig, J. Stat. Mech. Theor. Exp. **2013**, P01012 (2013).
[21] T. Held, A. Nourmohammad, and M. Lässig, J. Stat. Mech. Theor. Exp. **2014**, P09029 (2014).
[22] G. R. Price, Nature **227**, 520 (1970).
[23] R. Lande, Evolution **30**, 314 (1976).
[24] C. Gardiner, *Handbook of Stochastic methods: for physics, chemistry and the natural sciences* (Springer, 2004), 3rd ed.
[25] R. A. Fisher, *The genetical theory of natural selection* (Oxford University Press, USA, 1930), 1st ed.
[26] J. F. C. Kingman, J. Appl. Probab. **19**, 27 (1982).
[27] R. E. Bellman, *Dynamic Programming* (Princeton University Press, Princeton, NJ., 1957).
[28] R. E. Bellman, Proc. Natl. Acad. Sci. U.S.A. **40**, 231 (1954).
[29] H. J. Kappen, Phys. Rev. Lett. **95**, 200201 (2005).
[30] H. J. Kappen, J. Stat. Mech. Theor. Exp. **2005**, P11011 (2005).
[31] E. Todorov, in *Advances in Neural Information Processing Systems 19*, edited by B. Scholkopf, J. Platt, and T. Hoffman (MIT Press: Cambridge, MA, USA, 2007), pp. 1369–1376.
[32] J. L. Lush, Animal Breeding Plans (Iowa State Press, Ames, Iowa, 1937).
[33] R. Lande and S. J. Arnold, Evolution **37**, 1210 (1983).
[34] D. Tieman, G. Zhu, M. F. R. Resende, T. Lin, C. Nguyen, D. Bies, J. L. Rambla, K. S. O. Beltran, M. Taylor, B. Zhang, et al., Science **355**, 391 (2017).
[35] O. Rivoire and S. Leibler, J Stat Phys **142**, 1124 (2011).
[36] L. D. Landau and E. M. Lifshitz, *Statistical Physics; part 1* (Addison-Wesley, 1958), 1st ed.
[37] H. J. Kappen and H. C. Ruiz, J Stat Phys **162**, 1244 (2016).
[38] E. A. Theodorou and E. Todorov, in *Proceedings of 51st IEEE Conference on Decision and Control* (IEEE, Maui, HI, USA, 2012), pp. 1466–1473.
[39] W. Bialek and N. Tishby, arXiv (1999), cond-mat/9902341v1.
[40] R. E. Cobb, N. Sun, and H. Zhao, Methods **60**, 81 (2013).
[41] G. T. Hess, J. Tycko, D. Yao, and M. C. Bassik, Mol. Cell **68**, 26 (2017).
[42] K. Naran, T. Nundalall, S. Chetty, and S. Barth, Front Microbiol **9**, 3158 (2018).
[43] M. Łuksza and M. Lässig, Nature **507**, 57 (2014).
[44] R. A. Neher, C. A. Russell, and B. I. Shraiman, eLife **3**, e03568 (2014).
[45] J. Otwinowski and C. LaMont, arXiv (2019), 1912.03395.
[46] A. Nourmohammad, J. Rambeau, T. Held, V. Kovacova, J. Berg, and M. Lässig, Cell Rep. **20**, 1385 (2017).
[47] B. Wittenmark, in *Adaptive Systems in Control and Signal Processing 1995*, edited by C. Bányász (Oxford, UK, 1995), pp. 67–72.

Appendix A: Hamilton-Jacobi-Bellman equation for optimal control

We define a general stochastic evolutionary process for a population of size N with an evolutionary drive due to natural selection and mutations $A(\mathbf{x},t)$ and an external artificial selection $\mathbf{u}(\mathbf{x},t)$,

$$d\mathbf{x} = (A(\mathbf{x}) + \mathbf{u}(\mathbf{x},t))dt + \Sigma(\mathbf{x})d\mathbf{W} \tag{A1}$$

Here, time t is measured in units of the coalescence time N (i.e., the effective population size). $d\mathbf{W}$ is a differential random walk due to an underlying Wiener process with an amplitude Σ, which is the square root of the phenotypic covariance matrix K: $\Sigma^\top\Sigma \equiv K$. The stochastic evolution in eq. A1 defines an ensemble of phenotypic trajectories, the statistics of which can be characterized by a conditional probability density $P(\mathbf{x},t|\mathbf{x}',t')$ for a population to have a phenotype $\mathbf{x}$ at time t, given its state $\mathbf{x}'$ at a previous time $t' < t$. For a given artificial selection protocol $\mathbf{u}(\mathbf{x},t)$, the conditional probability density evolves according to a Fokker-Planck equation [1],

$$\frac{\partial}{\partial t}P(\mathbf{x},t|\mathbf{x}',t') = \left[\frac{1}{2}\mathbf{Tr}K\nabla_{\mathbf{xx}} - \nabla_{\mathbf{x}}\cdot(A(\mathbf{x}) + \mathbf{u}(\mathbf{x},t))\right]P(\mathbf{x},t|\mathbf{x}',t') \tag{A2}$$

where we have used the short hand notation, $\nabla_{\mathbf{x}}\cdot\mathcal{O} = \sum_i \frac{\partial}{\partial x_i}\mathcal{O}$ and $\mathbf{Tr}K\nabla_{\mathbf{xx}}\mathcal{O} = \sum_{i,j} K_{ij}\frac{\partial^2}{\partial x_i \partial x_j}\mathcal{O}$, as operators that act on the function $\mathcal{O}$ in front of them.

The purpose of artificial selection is to minimize a cost function,

$$\Omega(\mathbf{x},\mathbf{u},t) = V(\mathbf{x},t) + \frac{1}{2}\mathbf{u}^\top B\mathbf{u} \tag{A3}$$

where $V(\mathbf{x},t) \equiv V(|\mathbf{x}_t - \mathbf{x}^*|)$ is the cost for deviating from the desired target $\mathbf{x}^*$ during evolution and B is the cost for intervening in natural evolution and applying artificial selection $\mathbf{u} \equiv \mathbf{u}(\mathbf{x},t)$.

We define the cost-to-go function $J(\mathbf{x},t)$ as the expected value for the cumulative cost from time t to end of the process t_f, subject to the evolutionary dynamics and under an optimal control $\mathbf{u}^*_{t\to t_f}$,

$$J(\mathbf{x},t) = \min_{\mathbf{u}_{t\to t_f}}\left\langle Q(\mathbf{x},t_f) + \int_t^{t_f}\Omega(\mathbf{x}_s,\mathbf{u}_s)ds\right\rangle \tag{A4}$$

Here, $Q(\mathbf{x},t_f) \equiv Q(|\mathbf{x}_{t_f} - \mathbf{x}^*|)$ is the cost of deviation from the target at the end point t_f, which in general can be distinct from the path cost $V(\mathbf{x}_t)$. We can formulate a recursive relation for the cost-to-go function $J(\mathbf{x},t)$,

$$\begin{aligned}
J(\mathbf{x},t) &= \min_{\mathbf{u}_{t\to t_f}}\left\langle Q(\mathbf{x}_{t_f}) + \int_t^{t_f}\Omega(\mathbf{x}_s,\mathbf{u}_s)ds\right\rangle \\
&= \lim_{\delta t\to 0}\min_{\mathbf{u}_{t\to t_f}}\left\langle Q(\mathbf{x}_{t_f}) + \int_t^{t+\delta t}\Omega(\mathbf{x}_s,\mathbf{u}_s)ds + \int_{t+\delta t}^{t_f}\Omega(\mathbf{x}_s,\mathbf{u}_s)ds\right\rangle \\
&= \lim_{\delta t\to 0}\min_{\mathbf{u}_{t\to t_f}}\left\langle J(\mathbf{x}_{t+\delta t},t+\delta t) + \int_t^{t+\delta t}\Omega(\mathbf{x}_s,\mathbf{u}_s)ds\right\rangle \\
&= J(\mathbf{x}_t,t) + \min_{\mathbf{u}_{t\to t_f}}\left\langle \Omega(\mathbf{x}_s,\mathbf{u}_s)\delta t + \left[\frac{\partial}{\partial t}J(\mathbf{x}_t,t) + (A(\mathbf{x}_t)+\mathbf{u})^\top(\nabla J) + \frac{1}{2}\sum_{ij}K_{ij}\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}J\right]\delta t\right\rangle
\end{aligned} \tag{A5}$$

where we used Ito calculus to expand the cost-to-go function, $J(\mathbf{x}_{t+\delta t},t+\delta t)$; see e.g. ref. [2]. By reordering the terms in eq. A5, we arrive at the Hamilton-Jacobi-Bellman (HJB) equation,

$$\begin{aligned}
-\frac{\partial}{\partial t}J(\mathbf{x},t) &= \min_{\mathbf{u}}\left[\Omega(\mathbf{x}_t,\mathbf{u}_t) + A(\mathbf{x}_t)^\top\cdot\nabla J + \frac{1}{2}\mathbf{Tr}K\nabla_{\mathbf{xx}}J\right] \\
&= \min_{\mathbf{u}}\left[\frac{1}{2}\mathbf{u}^\top B\mathbf{u} + \mathbf{u}^\top\cdot\nabla J\right] + V(\mathbf{x}) + (A(\mathbf{x}_t)+\mathbf{u})^\top\cdot\nabla J + \frac{1}{2}\mathbf{Tr}K\nabla_{\mathbf{xx}}J
\end{aligned} \tag{A6}$$

The functional form for the optimal artificial selection $\mathbf{u}^*$ follows by minimizing the right hand side of eq. A6 with respect to $\mathbf{u}$,

$$\mathbf{u}^* = -B^{-1}\nabla J. \tag{A7}$$

Therefore, the time- and phenotype-dependent solution for the cost-to-go function $J(\mathbf{x},t)$ determines the optimal protocol for artificial selection $\mathbf{u}^*(\mathbf{x},t)$. By substituting the form of the optimal control $\mathbf{u}^*$ in eq. A6, we arrive at a non-linear partial differential equation for the cost-to-go function,

$$-\frac{\partial}{\partial t}J(\mathbf{x},t) \;=\; -\frac{1}{2}(\nabla J)^\top B^{-1}\nabla J + V(\mathbf{x}) + A(\mathbf{x}_t)^\top\cdot\nabla J + \frac{1}{2}\mathbf{Tr}K\nabla_{\mathbf{xx}}J \tag{A8}$$

which should be solved with a boundary condition $J(\mathbf{x},t_f) = Q(\mathbf{x},t_f)$ at the end point. We introduce a new variable $\Psi = \exp[-J/\lambda]$ as the exponential of the cost-to-go function. The dynamics of Ψ follows,

$$\frac{\lambda}{\Psi}\frac{\partial}{\partial t}\Psi \;=\; -\frac{\lambda^2}{2\Psi^2}(\nabla\Psi)^\top B^{-1}\nabla\Psi + V(\mathbf{x}) - \frac{\lambda}{\Psi}A(\mathbf{x}_t)^\top\cdot(\nabla\Psi) - \frac{\lambda}{2}K\left[\frac{-1}{\Psi^2}(\nabla\Psi)^\top\cdot\nabla\Psi + \frac{1}{\Psi}\nabla_{\mathbf{xx}}\Psi\right] \tag{A9}$$

The dynamics of Ψ linearizes if and only if there exists a scalar λ that relates the control cost to the covariance matrix such that $B = \lambda K^{-1}$. This criteria is known as the path-integral control condition [3, 4] by which we can map a generally non-linear control problem onto a linear stochastic process. The path-integral control condition implies that the cost of artificial selection on each phenotype should be inversely proportional to the phenotype's fluctuations. In other words, artificially tweaking with highly conserved phenotypes should be more costly than with variable phenotypes. In this case, the HJB equation for the transformed cost-to-go function Ψ follows,

$$\frac{\partial}{\partial t}\Psi = -A(\mathbf{x})^\top\cdot\nabla\Psi - \frac{1}{2}\mathbf{Tr}K\nabla_{\mathbf{xx}}\Psi + \frac{1}{\lambda}V(\mathbf{x})\Psi \equiv -L^\dagger\Psi \tag{A10}$$

where $L^\dagger$ is a linear operator acting on the function Ψ. Equation A10 has the form of a backward Fokker-Planck equation with the boundary condition at the end point $\Psi(\mathbf{x},t_f) = \exp[-J(\mathbf{x},t_f)/\lambda] = \exp[Q(\mathbf{x},t_f)/\lambda]$. We can define a conjugate function P_u that evolves forward in time according to the Hermitian conjugate of the operator $L^\dagger$. This conjugate operator L can be characterized by evaluating the inner product of the two functions,

$$\begin{aligned}
\langle LP_u|\Psi\rangle = \left\langle P_u|L^\dagger\Psi\right\rangle \;&=\; \int d\mathbf{x}\, P_u(\mathbf{x},t)L^\dagger\Psi(\mathbf{x},t)\\
&=\; \int d\mathbf{x}\, P_u(\mathbf{x},t)\left(A(\mathbf{x})^\top\cdot\nabla\Psi + \frac{1}{2}\mathbf{Tr}K\nabla_{\mathbf{xx}}\Psi - \frac{1}{\lambda}V(\mathbf{x})\Psi\right)\Psi(\mathbf{x},t)\\
&=\; \int d\mathbf{x}\left(-\frac{1}{\lambda}V(\mathbf{x})P_u(\mathbf{x},t) - \nabla A(\mathbf{x})P_u + \frac{1}{2}\mathbf{Tr}K\nabla_{\mathbf{xx}}P_u\right)^\top\Psi(\mathbf{x},t)
\end{aligned}$$

where we performed integration by part and assumed that the function P_u vanishes at the boundaries. This formulation suggests a forward evolution by the operator $L^\dagger$ for the function $P_u(\mathbf{x}',t'|\mathbf{x},t)$,

$$\frac{\partial}{\partial t}P_u(\mathbf{x}',t'|\mathbf{x},t) = LP_u(\mathbf{x}',t'|\mathbf{x},t) = \left[\underbrace{\frac{1}{2}\mathbf{Tr}K\nabla_{\mathbf{xx}} - \nabla A(\mathbf{x})}_{L_0} - \frac{1}{\lambda}V(\mathbf{x})\right]P_u(\mathbf{x}',t'|\mathbf{x},t). \tag{A11}$$

with a boundary condition with the initial time point $P(\mathbf{x}',t|\mathbf{x},t) = \delta(\mathbf{x}-\mathbf{x}')$. Importantly, the linear operator L resembles the forward Fokker Planck operator L_0 for evolution under natural selection (i.e., the dynamics in eq. A2 with $\mathbf{u}=0$) with an extra annihilation term $V(\mathbf{x})/\lambda$. The evolutionary dynamics with the L_0 operator under natural selection conserves the probability density. The annihilation term on the other hand, eliminates the evolutionary trajectories with a rate proportional to their cost $V(\mathbf{x},t)/\lambda$ at each time point.

Since Ψ evolves backward in time according to $L^\dagger$ and P_u evolves forward in time according to L, their inner product $\langle P_u|\Psi\rangle = \int d\mathbf{x}' P_u(\mathbf{x}',t'|\mathbf{x},t)\Psi(\mathbf{x}',t')$ remains time-invariant[1]. Therefore, the inner product of the two functions at

---

[1] The inner product of the two conjugate functions $\langle P_u|\Psi\rangle$ is time invariant:

$$\begin{aligned}
\left\langle P_u(t')|\Psi(t')\right\rangle \equiv \int d\mathbf{x}' P_u(\mathbf{x}',t'|\mathbf{x},t)\Psi(\mathbf{x}',t') \;&=\; \left\langle e^{L(t'-t)}P_u(t)|e^{-L^\dagger(t'-t)}\Psi(t)\right\rangle\\
&=\; \left\langle P_u(t)|e^{L^\dagger(t'-t)}e^{-L^\dagger(t'-t)}\Psi(t)\right\rangle \equiv \langle P_u(t)|\Psi(t)\rangle
\end{aligned}$$

the initial and the final time points are equal, which follows,

$$\langle P_u(t)|\Psi(t)\rangle = \langle P_u(t_f)|\Psi(t_f)\rangle \rightarrow \quad \int d\mathbf{x}' P_u(\mathbf{x}', t|x, t)\Psi(\mathbf{x}', t) = \int d\mathbf{x}' P_u(\mathbf{x}', t_f|x, t)\Psi(\mathbf{x}', t_f) \tag{A12}$$

$$\rightarrow \quad \Psi(\mathbf{x}, t) = \int d\mathbf{x}' P_u(\mathbf{x}', t_f|\mathbf{x}, t)\exp[-Q(\mathbf{x}', t_f)/\lambda] \tag{A13}$$

where we substituted the boundary condition for P_u at the initial time t and for Ψ at the finial time t_f. Thus, the cost-to-go function follows,

$$J(\mathbf{x}, t) = -\lambda \log \Psi(\mathbf{x}, t) \;=\; -\lambda \log \int d\mathbf{x}' P_u(\mathbf{x}', t_f|\mathbf{x}, t)\exp[-Q(\mathbf{x}', t_f)/\lambda] \tag{A14}$$

Appendix B: Path integral solution to stochastic adaptive control

Given the structure of the linear forward operator L, we can either exactly compute the conditional function $P_u(\mathbf{x}', t_f|\mathbf{x}, t)$ or to use approximation methods common for path integrals (e.g. the semi-classical method) to evaluate cost-to-go function in eq. A14. To formulate a path integral for $P_u(\mathbf{x}', t_f|\mathbf{x}, t)$, we discretize the time window $[t : t_f]$ into n small time slices of length ϵ, $(t_0, t_1, \ldots, t_n)$, with $n\epsilon = t_f - t$. The conditional probability $P_u(\mathbf{x}', t_f|\mathbf{x}, t)$ can be written as an integral over all trajectories that start at the phenotypic state $\mathbf{x}$ at time $t_0 \equiv t$ and end at $\mathbf{x}'$ at time $t_n \equiv t_f$,

$$P_u(\mathbf{x}', t_f|\mathbf{x}, t) \sim \int \prod_{i=1}^{n} d\mathbf{x}_i \, P_u(\mathbf{x}_i, t_i|\mathbf{x}_{i-1}, t_{i-1})\, \delta(\mathbf{x}_n - \mathbf{x}') \tag{B1}$$

The short-time propagator $P_u(\mathbf{x}_i, t_i|\mathbf{x}_{i-1}, t_{i-1})$ follows a simple Gaussian form [1],

$$\begin{aligned} P_u(\mathbf{x}_i, t_i|\mathbf{x}_{i-1}, t_{i-1}) \;&\sim\; \exp\left\{-\frac{1}{\lambda}\left[\Big(\mathbf{x}_i - \mathbf{x}_{i-1} - \epsilon A(\mathbf{x}_i)\Big)^\top \frac{\lambda K^{-1}}{2\epsilon}\Big(\mathbf{x}_i - \mathbf{x}_{i-1} - \epsilon A(\mathbf{x}_i)\Big) + V(\mathbf{x}_i)\epsilon\right]\right\} \\ &=\; \exp\left\{-\frac{\epsilon}{\lambda}\left[\left(\frac{\mathbf{x}_i - \mathbf{x}_{i-1}}{\epsilon} - A(\mathbf{x}_i)\right)^\top \frac{B}{2}\left(\frac{\mathbf{x}_i - \mathbf{x}_{i-1}}{\epsilon} - A(\mathbf{x}_i)\right) + V(\mathbf{x}_i)\right]\right\} \end{aligned} \tag{B2}$$

where we used, $K = \lambda B^{-1}$. We can express the cost-to-go function (eq. A14) as a path integral,

$$\begin{aligned} e^{-J(\mathbf{x},t)/\lambda} \;&=\; \int d\mathbf{x}' P_u(\mathbf{x}', t_f|\mathbf{x}, t)\exp[-Q(\mathbf{x}', t_f)/\lambda] \\ &\sim\; \int \prod_{i=1}^{n} d\mathbf{x}_i \exp\left\{-\frac{\epsilon}{\lambda}\left[\left(\frac{\mathbf{x}_i - \mathbf{x}_{i-1}}{\epsilon} - A(\mathbf{x}_i)\right)^\top \frac{B}{2}\left(\frac{\mathbf{x}_i - \mathbf{x}_{i-1}}{\epsilon} - A(\mathbf{x}_i)\right) + V(\mathbf{x}_i) + \frac{Q(\mathbf{x}_n)}{\epsilon}\right]\right\} \\ &\sim\; \int \mathcal{D}(\mathbf{x}) \exp\left[-\frac{1}{\lambda}\left(Q(\mathbf{x}(t_f)) + \int_t^{t_f} dt \left[\left(\frac{d\mathbf{x}(t)}{dt} - A(\mathbf{x}(t), t)\right)^\top \frac{B}{2}\left(\frac{d\mathbf{x}(t)}{dt} - A(\mathbf{x}(t), t)\right) + V(\mathbf{x}, t)\right]\right)\right] \\ &\equiv\; \int \mathcal{D}(\mathbf{x}) \exp\left[-\frac{1}{\lambda} S_{\text{path}}(\mathbf{x}(t \rightarrow t_f))\right] \end{aligned} \tag{B3}$$

where $S_{\text{path}}(\mathbf{x}(t \rightarrow t_f))$ is a corresponding action and $\mathcal{D}(\mathbf{x}) \sim \prod_{i=1}^{n} d\mathbf{x}_i$ is the integral measure over all the trajectories that start at $\mathbf{x}_0 = \mathbf{x}(t)$. Numerically, this formulation provides a way to generate evolutionary trajectories under artificial selection as an exponentially weighted ensemble from trajectories under natural selection [3, 4]. Moreover, if λ is small, the integral is dominated by the trajectories that are close to the most likely (classical) trajectory $\hat{\mathbf{x}}(t \rightarrow t_f)$, and the path integral can be approximated using the semi-classical method; see ref. [3].

Appendix C: End point control with quadratic cost

In the case that the path cost is zero $V(\mathbf{x}) = 0$, the artificially and naturally selected trajectories become distinct only due to the contributions from the end-point cost at $t = t_f$. For the choice of a linear evolutionary force $A(\mathbf{x}) = -2KC\mathbf{x}$ and a quadratic end-point cost, $Q(\mathbf{x}, t_f) = \frac{1}{2}(\mathbf{x}_{tf} - \mathbf{x}^*)^\top \tilde{G}(\mathbf{x}_{tf} - \mathbf{x}^*)$, evolution follows an Ornstein-Uhlenbeck (OU) process and the solution to eq. (A11) takes a Gaussian form (see e.g. ref. [2]),

$$P_u(\mathbf{x}_{t_f}, t_f|\mathbf{x}, t) \sim \exp\left[\frac{-1}{2}\left(\mathbf{x}_{t_f} - \mu(\mathbf{x}, \tau)\right)^\top K_\tau^{-1}\left(\mathbf{x}_{t_f} - \mu(\mathbf{x}, \tau)\right)\right] \tag{C1}$$

with a conditional time-dependent mean,

$$\mu(\mathbf{x}, \tau) \;=\; \exp[-2KC\tau]\mathbf{x}, \tag{C2}$$

and a covariance matrix,

$$K_\tau \;=\; \int_t^{t_f} dt' \exp[-2KC(t_f - t')]K \exp[-2CK(t_f - t')]. \tag{C3}$$

with $\tau = t_f - t$.

In this case, the cost-to-go in eq. (A14) can be evaluated by a Gaussian integral to marginalize over the end state $\mathbf{x}_{t_f}$,

$$\begin{aligned}\exp[-J/\lambda] \;&\sim\; \int d\mathbf{x}_{t_f} \exp\left[\frac{-1}{2\lambda}(\mathbf{x}_{t_f} - \mu(\mathbf{x}, \tau))^\top \lambda K_\tau^{-1}(\mathbf{x}_{t_f} - \mu(\mathbf{x}, \tau))\right] \exp\left[\frac{-1}{2\lambda}(\mathbf{x}_{t_f} - \mathbf{x}^*)^\top \tilde{G}(\mathbf{x}_{t_f} - \mathbf{x}^*)\right] \\ &\sim\; \exp\left[\frac{1}{2\lambda}(\mu(\mathbf{x}, \tau) - \mathbf{x}^*)^\top \left(\tilde{G}\left[\lambda K_\tau^{-1} + \tilde{G}\right]^{-1}\tilde{G} - \tilde{G}\right)(\mu(\mathbf{x}, \tau) - \mathbf{x}^*)\right]\end{aligned} \tag{C4}$$

resulting in an optimal artificial selection protocol,

$$\begin{aligned}\mathbf{u}^* = -B^{-1}\nabla J \;&=\; -\frac{K}{\lambda}\nabla J = -\frac{K}{\lambda}\left[\nabla\mu(\mathbf{x}, \tau)\right]^\top \left[\tilde{G} - \tilde{G}\left[\lambda K_\tau^{-1} + \tilde{G}\right]^{-1}\tilde{G}\right](\mu(\mathbf{x}, \tau) - \mathbf{x}^*) \\ &=\; -\frac{K}{\lambda}\exp[-2CK\tau]\tilde{G}\left[I - \frac{K_\tau}{\lambda}\left[I + \frac{K_\tau}{\lambda}\tilde{G}\right]^{-1}\tilde{G}\right](e^{-2KC\tau}\mathbf{x} - \mathbf{x}^*)\end{aligned} \tag{C5}$$

As the time approaches the end point ($t \to t_f$ or $\tau \to 0$), optimal artificial selection acts as a linear attractive force (i.e., a stabilizing selection)

$$\mathbf{u}^*(\tau \to 0) = \frac{-1}{\lambda}K\tilde{G} \cdot (\mathbf{x} - \mathbf{x}^*) + \mathcal{O}(\tau) \tag{C6}$$

to maintain the population close to the phenotypic target, with an effective strength of artificial stabilizing selection $\tilde{G}/\lambda$.

Appendix D: Artificial selection with intermittent monitoring

Here, we assume that artificial selection is imposed in discrete steps with time interval τ. Similar to the continuous control, the cost function has two components: the cost of control at the end of each intervention and the cumulative cost of deviation from the optimum throughout each interval. The stationary cost-to-go function follows,

$$\begin{aligned}&J(\mathbf{x}, t_m; \tau) \\ &= \min_{\mathbf{u}} \lim_{M\to\infty} \frac{1}{(M - m)\tau}\left\langle \sum_{i=m}^{M} \mathbf{u}_i^\top B\mathbf{u}_i + \int_{t_i}^{t_M} V(\mathbf{x}_t)dt\right\rangle_{\text{evol.}}\end{aligned} \tag{D1}$$

where we have normalized the path cost by the interval τ to assure that the cost-to-go remains finite.

To further simplify, we only consider one dimensional phenotype x with intra-population variance k, the cost of deviation $V(x) = g(x - x^*)^2/2$ from target x^*, and the cost of intervention $\beta u^2/2$ with artificial selection u. In the stationary state and in the regime of small interventions ($gk/\lambda < 1$), we assume that the optimal artificial selection protocol u^* should be a variant of the case with full information with a strength of selection α_τ dependent on the time window τ, $u^*_\tau = -k\alpha_\tau(x - x^*)$. Our goal is to characterize the strength of optimal artificial selection α_τ.

The total cost over an interval τ in the stationary state follows,

$$\begin{aligned}\Omega_\tau(x) &= \frac{\beta}{2}\langle u^2\rangle + \frac{1}{\tau}\left\langle \int_{t=t_i}^{t_i+\tau} V(x_t)dt \right\rangle \\ &= \left\langle \frac{\beta}{2}k^2\alpha_\tau^2(x_\tau - x^*)^2 + \frac{1}{2\tau}\gamma\int_{t=0}^{\tau}(x_t - x^*)^2 dt \right\rangle\end{aligned} \tag{D2}$$

We are interested in the regime of moderate to weak interventions ($gk/\lambda < 1$), for which the linear response theory can characterize the change in the state of the system after each intervention. In this regime, evolution under artificial selection can be approximated as a perturbation from the dynamics under natural selection. The evolutionary dynamics of the phenotype distribution is governed by a time-dependent Fokker Planck operator, $L(x,t)$,

$$\frac{\partial}{\partial t}P_{u;\tau}(x,t) = [L_0(\mathbf{x}) + L_u(\mathbf{x})Y(t)]P_{u;\tau}(x,t) \tag{D3}$$

where $P_{u;\tau}(\mathbf{x},t)$ is the full probability density under intermittent artificial selection, which can be split into the stationary solution under natural selection and the time-dependent deviation due to artificial selection: $P_{u;\tau}(x,t) = P_0(\mathbf{x}) + P_u(\mathbf{x},t;\tau)$. $L_0(\mathbf{x})$ is the Fokker Planck operator associated with evolution under natural selection (i.e., the dynamics in eq. A2 with $\mathbf{u} = 0$), $L_u(\mathbf{x}) = \partial_x k\alpha_\tau(x - x^*)$ is the state-dependent operator associated with artificial selection and $Y(t) = \lim_{M\to\infty}\sum_{i=1}^{M}\delta(t - t_i)$ characterizes a time-dependent component due to artificial selection interventions at the end of each time interval.

In the regime of linear response, where the impact of artificial selection is small, the deviation $\langle\Delta z\rangle$ of an expected value of an arbitrary function $\langle z(\mathbf{x})\rangle$ from it stationary state (i.e., under natural selection) follows (cf. [1]),

$$\langle\Delta z(t)\rangle = \int z(\mathbf{x})P_u(x,t)d\mathbf{x} \equiv \int_{-\infty}^{\infty} R_{z,L_u}(t - t')Y(t')dt' \tag{D4}$$

where $R_{z,L_u}(t)$ is the response function to artificial selection L_u,

$$R_{z,L_u}(t) = \begin{cases}\int z(\mathbf{x})\, e^{L_0(\mathbf{x})\cdot t}\, L_u(\mathbf{x})P_0(\mathbf{x})d\mathbf{x} & \text{for } t \geq 0 \\ 0 & \text{for } t < 0\end{cases} \tag{D5}$$

At end of each time interval, artificial selection imposes a shock-type perturbation to the evolutionary process. The immediate response of the population to this selection pressure can be characterized by the instantaneous response function (i.e., with $Y(t') = \delta(t - t')$), resulting in the change in a given phenotypic statistics z (cf. [1]),

$$\langle\Delta z(t)\rangle = \int z(x)L_u(\mathbf{x})P_0(x)dx \tag{D6}$$

$$= \frac{1}{Z}\int dx\, z(x)\,\frac{\partial}{\partial x}\left(k\alpha_\tau\,(x - x^*)\,\exp\left[-\frac{x^2}{2\mathrm{var}_{\mathrm{st}_0}}\right]\right) \tag{D7}$$

$$= k\alpha_\tau\left(\langle z(x)\rangle_{\mathrm{st}_0} - \frac{1}{\mathrm{var}_{\mathrm{st}_0}}\left[\left\langle z(x)\,x^2\right\rangle_{\mathrm{st}_0} - x^*\langle z(x)\,x\rangle_{\mathrm{st}_0}\right]\right) \tag{D8}$$

where $P_0(x)$ is the Gaussian stationary distribution for the mean phenotype under natural selection, $\mathrm{var}_{\mathrm{st}_0} = 1/4c$ is the stationary ensemble variance for the mean phenotype under natural selection, Z is the normalization factor for the underlying stationary distribution, and $\langle\cdot\rangle_{\mathrm{st}_0}$ indicates expectation values under the stationary distribution.

At the beginning of each interval $t = 0$ the deviation of the mean phenotype from its expectation under natural selection $\langle x\rangle_{\mathrm{st}_0} = 0$ follows,

$$\langle\Delta x\rangle = \langle x(t = 0)\rangle = k\alpha_\tau x^* \tag{D9}$$

Similarly, the deviation in the second moment of the phenotypic distribution from the expectation under natural selection $\langle x^2\rangle_{\rm st_0} = {\rm var}_{\rm st_0}$ follows,

$$\langle \Delta x^2\rangle = \langle x^2(t=0)\rangle - {\rm var}_{\rm st_0} = \alpha_\tau \left({\rm var}_{\rm st_0} - \frac{1}{{\rm var}_{\rm st_0}}\langle x^4\rangle_{\rm st_0}\right) = -2k\alpha_\tau\, {\rm var}_{\rm st_0} = -k\alpha_\tau/2c \tag{D10}$$

Thus, the phenotypic variance at the beginning of each interval follows,

$${\rm var}_u(t=0) = \langle x^2(t=0)\rangle - \langle x(t=0)\rangle^2 = {\rm var}_{\rm st_0}\left[1-2k\alpha_\tau\right] - \left[k\alpha_\tau\, x^*\right]^2 \tag{D11}$$

Following an intervention at time $t=0$, populations evolve according to natural selection until the next intervention. Therefore, the phenotype statistics during each time interval ($0<t<\tau$) follow,

$$\langle x(t)\rangle \;=\; \alpha_\tau x^* e^{-2kct} \tag{D12}$$

$${\rm var}(t) \;=\; \left[{\rm var}_{\rm st_0}(1-4kc) - (k\alpha_\tau\, x^*)^2\right] e^{-4kct} + {\rm var}_{\rm st_0}(1-e^{-4kct}) = {\rm var}_{\rm st_0}(1-2k\alpha_\tau\, e^{-4kct}) - (k\alpha_\tau\, x^*)^2\, e^{-4kct} \tag{D13}$$

We can now evaluate the cumulative cost function (eq. D2)

$$\begin{aligned}
\Omega_\tau(x) \;&=\; \frac{\beta}{2}\langle u^2\rangle + \frac{1}{2\tau}\gamma\left\langle \int_{t=0}^{\tau}(x_t-x^*)^2 dt\right\rangle \\
&=\; \left\langle \frac{\beta}{2}k^2\alpha_\tau^2(x_\tau-x^*)^2 + \frac{1}{2\tau}\gamma\int_{t=0}^{\tau}(x_t-x^*)^2 dt\right\rangle \\
&=\; \frac{\beta}{2}k^2\alpha_\tau^2\left[(\langle x_\tau\rangle - x^*)^2 + {\rm var}(\tau)\right] + \frac{1}{2\tau}\gamma\int_{t=0}^{\tau}\left[(\langle x_t\rangle - x^*)^2 + \sigma^2(t)\right] dt \\
&=\; \frac{\beta}{2}k^2\alpha_\tau^2\left[{\rm var}_{\rm st_0}(1-2k\alpha_\tau e^{-4kc\tau}) + (x^*)^2(1-2k\alpha_\tau e^{-2kc\tau})\right] \\
&\quad + \frac{1}{2\tau}\gamma\left[\frac{-\alpha_\tau}{2c}\left((1-e^{-4kc\tau}){\rm var}_{\rm st_0} + 2(1-e^{-2kc\tau})(x^*)^2\right) + ({\rm var}_{\rm st_0} + (x^*)^2)\tau\right]
\end{aligned} \tag{D14}$$

where we have used the time-dependent expectation and variance in eqs. D12 and D13.

The optimal strength of artificial selection α_τ^* for intermittent interventions can be characterized by minimizing the cost function (eq. D14) with respect to α_τ,

$$\alpha_\tau^* \;=\; \frac{\gamma}{\lambda}\left(\frac{(1-e^{-\tau})(1+8c(x^*)^2+e^{-\tau})}{2\tau(1+4c(x^*)^2)}\right) + \mathcal{O}\left((k\gamma/\lambda)^2\right) \tag{D15}$$

which in the limit of small separation time ($\tau\to 0$) approaches the expectation under continuous monitoring in the stationary state (eq. C6), $\alpha^*(\tau\to 0) = \gamma/\lambda$.

Appendix E: Work performed by artificial selection

Artificial selection changes the distribution of phenotypic trajectories $\mathbf{x}_{t_0}^{t_f} \equiv (\mathbf{x}_{t_0},\ldots,\mathbf{x}_{t_f})$ from $P_0(\mathbf{x}_{t_0}^{t_f})$ in the stationary state under natural selection to a configuration closer to the desired target $P_{\mathbf{u}}(\mathbf{x}_{t_0}^{t_f})$. In analogy to thermodynamics, we can associate a free energy to these distributions, as $F_0 = \log P_0(\mathbf{x}_{t_0}^{t_f})$ and $F_u = \log P_u(\mathbf{x}_{t_0}^{t_f})$ (cf. [5]). The expected difference between the free energy of the two phenotypic configurations can be interpreted as the amount of work done by artificial selection, which corresponds to the Kullback-Leibler distance between the two distributions,

$$W_u \equiv \langle F_u\rangle - \langle F_0\rangle = \int \mathcal{D}\mathbf{x}\, P_u(\mathbf{x}_{t_0}^{t_f})\log\left[\frac{P_{\mathbf{u}}(\mathbf{x}_{t_0}^{t_f})}{P_0(\mathbf{x}_{t_0}^{t_f})}\right] \equiv D_{KL}(P_{\mathbf{u}}(\mathbf{x}_{t_0}^{t_f})||P_0(\mathbf{x}_{t_0}^{t_f})) \tag{E1}$$

where $\mathcal{D}\mathbf{x}$ is the integral measure over all trajectories. The estimate of work in eq. E1 however should not be interpreted as a physical work, rather as an information theoretical measure of discrimination between the two phenotype distributions due to artificial selection.

The evolution of the distribution for phenotype trajectories $P_u(\mathbf{x}_{t_0}^{t_f})$ under a given artificial selection protocol, $\mathbf{u}_{t_0}^{t_f}$ is Markovian (eqs. A1,A2). To characterize this path probability density, we will follow the path integral formulation in eq. B1 and discretize the time window $[t_0 : t_f]$ into n small time slices of length ϵ, $(t_0, t_1, \ldots, t_n)$, with $n\epsilon = t_f - t$. The probability of a given trajectory $P_u(\mathbf{x}_{t_0}^{t_f})$ can be written as a product of short-term propagators (i.e., conditional probabilities); cf. [1],

$$
\begin{aligned}
P_u(\mathbf{x}_{t_0}^{t_f}) &= \lim_{\epsilon\to 0} \prod_{s=t_0}^{t_f} P_u(\mathbf{x}_{s+\epsilon}|\mathbf{x}_s)) && \text{(E2)}\\
&= \lim_{n\to\infty} \prod_{i=1}^{n} \frac{1}{Z_i} \exp\left[-(\mathbf{x}_{i+1} - \mathbf{x}_i - \epsilon(A(\mathbf{x}_i) + \mathbf{u}(\mathbf{x}_i)))^\top \frac{K^{-1}}{2\epsilon}(\mathbf{x}_{i+1} - \mathbf{x}_i - \epsilon(A(\mathbf{x}_i) + \mathbf{u}(\mathbf{x}_i)))\right] && \text{(E3)}\\
&\sim P_0(\mathbf{x}_{t_0}^{t_f}) \lim_{n\to\infty} \prod_{i=1}^{n} e^{\mathbf{u}^\top(\mathbf{x}_i)K^{-1}(\mathbf{x}_{i+1}-\mathbf{x}_i-\epsilon A(\mathbf{x}_i))} \times e^{-\frac{\epsilon}{2}\mathbf{u}(\mathbf{x}_i)^\top K^{-1}\mathbf{u}(\mathbf{x}_i)} && \text{(E4)}\\
&= P_0(\mathbf{x}_{t_0}^{t_f}) \exp\left[-\int_{t_0}^{t_f} dt \frac{1}{2}\mathbf{u}^\top(\mathbf{x},t)K^{-1}\mathbf{u}(\mathbf{x},t) + \int_{t_0}^{t_f} \mathbf{u}^\top(\mathbf{x},t)K^{-1}(d\mathbf{x}_t - A(\mathbf{x}_t)dt)\right] && \text{(E5)}
\end{aligned}
$$

The Kullback-Leibler distance between the two distributions follows,

$$
\begin{aligned}
D_{KL}(P_{\mathbf{u}}(\mathbf{x}_{t_0}^{t_f})||P_0(\mathbf{x}_{t_0}^{t_f})) &= \int \mathcal{D}\mathbf{x}\, P_{\mathbf{u}}(\mathbf{x}_{t_0}^{t_f}) \left[-\int_{t_0}^{t_f} dt \left(\frac{1}{2}\mathbf{u}^\top(\mathbf{x},t)K^{-1}u(\mathbf{x},t)\right) + \int_{t_0}^{t_f} \mathbf{u}^\top(\mathbf{x},t)K^{-1}(d\mathbf{x}_t - A(\mathbf{x}_t)dt)\right]\\
&= \int_{t_0}^{t_f} dt \int \mathcal{D}\mathbf{x}\, P_{\mathbf{u}}(\mathbf{x}_{t_0}^{t_f}) \left(\frac{1}{2}\mathbf{u}^\top(\mathbf{x},t)K^{-1}\mathbf{u}(\mathbf{x},t)\right) \equiv \left\langle \frac{1}{2}\left(\mathbf{u}_{t_0}^{t_f}\right)^\top K^{-1}\mathbf{u}_{t_0}^{t_f}\right\rangle && \text{(E6)}
\end{aligned}
$$

where we have used $d\mathbf{x}_t - A(\mathbf{x}_t)dt = \mathbf{u}(\mathbf{x}_t,t)dt + d\mathbf{W}_t$, with $d\mathbf{W}_t$ as the stochastic differential measure for a multivariate Wiener process (cf. [2]). Importantly, with the criteria of path integral control (i.e., $K^{-1} = B/\lambda$), the Kullback-Leibler distance between the artificially and naturally selected phenotype distributions is equivalent to the cumulative cost of intervention, divided by the overall cost of artificial selection λ,

$$
D_{KL}(P_{\mathbf{u}}(\mathbf{x}_{t_0}^{t_f})||P_0(\mathbf{x}_{t_0}^{t_f})) = \frac{1}{2\lambda}\left\langle \left(\mathbf{u}_{t_0}^{t_f}\right)^\top B\, \mathbf{u}_{t_0}^{t_f}\right\rangle \qquad \text{(E7)}
$$

which can intuitively be interpreted as the amount of work done by artificial selection.

[1] H. Risken, C. Schmid, and W. Weidlich, Z. Physik **194**, 337 (1966).
[2] C. Gardiner, *Handbook of Stochastic methods: for physics, chemistry and the natural sciences* (Springer, 2004), 3rd ed.
[3] H. J. Kappen, Phys. Rev. Lett. **95**, 200201 (2005).
[4] H. J. Kappen, J. Stat. Mech. Theor. Exp. **2005**, P11011 (2005).
[5] L. D. Landau and E. M. Lifshitz, *Statistical Physics; part 1* (Addison-Wesley, 1958), 1st ed.

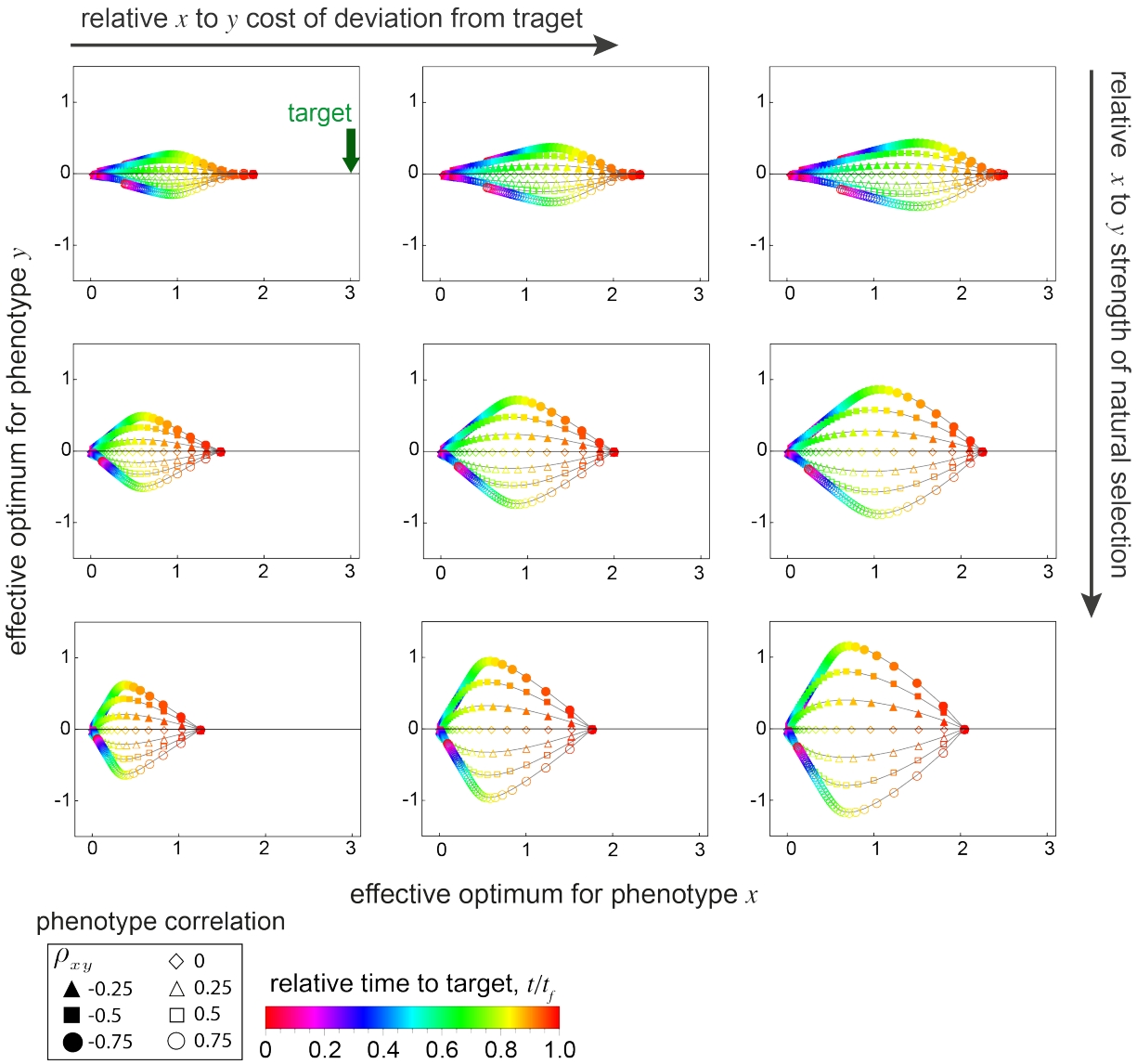

FIG. S1: **Effective fitness optimum for 2D covarying phenotypes under artificial selection.** The dynamics of the effective fitness peak is shown over time (colors) for 2D covarying phenotypes with correlations ρ_{xy} indicated by the shape of the markers, and for increasing end-point cost of deviation from target along the x-phenotype, $g_x = 1, 2, 3$ from left to right, with $g_y = 2$ and for increasing strength of natural selection on the x-phenotype, $c_x = 3, 5, 7$ from top to bottom with $c_y = 5$. Other parameters: $x^* = 3$, $y^* = 0$; $c_{xy} = 0$; $k_x = k_y = 0.02$; $\lambda = 0.1$.

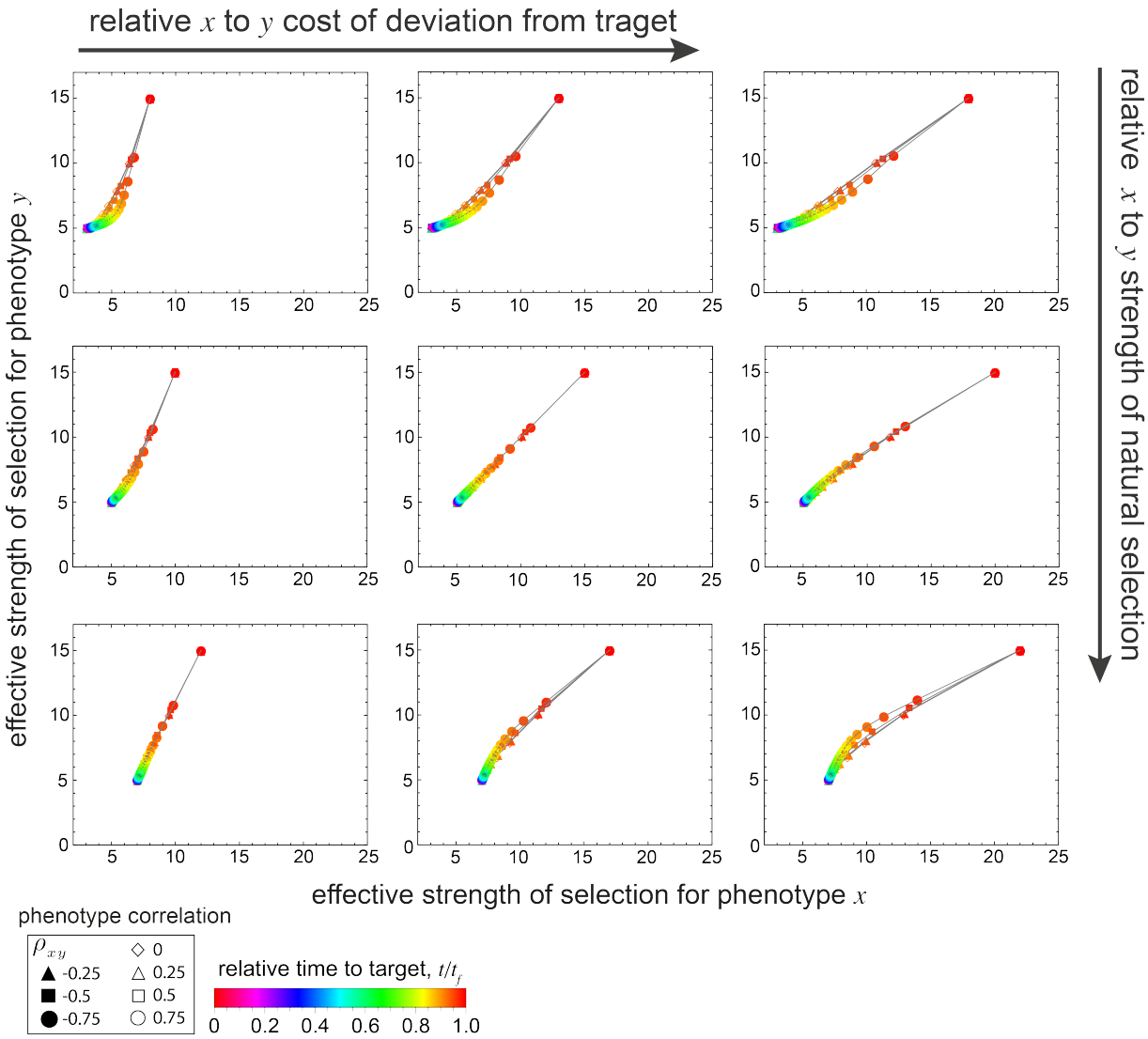

FIG. S2: **Effective strength of selection for 2D covarying phenotypes under artificial selection .** The dynamics of the effective strength of selection is shown over time (colors) for 2D covarying phenotypes with correlations ρ_{xy} indicated by the shape of the markers. The parameters in each panels are the same as in Fig. S1.