Forecasting Evolution: Mining for Predictability in the Unpredictable

Philip J. Gerrish, Ramiro Dominguez Aguilar, and Alexandre Colato

Contents

P. J. Gerrish (✉)
Department of Mathematics and Physics, Universidad Autónoma del Estado de Hidalgo, Pachuca de Soto, Mexico

Department of Biology, University of New Mexico, Albuquerque, NM, USA
e-mail: pgerrish@unm.edu

R. D. Aguilar
Department of Mathematics and Physics, Universidad Autónoma del Estado de Hidalgo, Pachuca de Soto, Mexico

A. Colato
Department of Natural Sciences, Mathematics and Education, Federal University of São Carlos, Araras, SP, Brazil

B. Sriraman (ed.), *Handbook of Visual, Experimental and Computational Mathematics*,
https://doi.org/10.1007/978-3-032-16368-4_12

Abstract

Biological evolution is an enormously high-dimensional process fraught with contingencies and nonlinearities. For this reason, evolutionary biology has traditionally been considered an exclusively retrospective science. This state of affairs, however, is beginning to change, and notions of predicting evolution are increasingly being discussed with straight faces – a cultural shift brought about by high-throughput and high-resolution biotech, ever-increasing computational capacity, the explosion of machine learning and AI, as well as the increasing need for evolutionary prediction in our rapidly changing world. Previous assessments of predictability in evolution have, wittingly or not, employed something like a *frequentist* approach, whereby a fitness landscape and its derivative distribution of fitness effects (DFEs) are determined beforehand through enormously expensive and time-consuming lab work, and evolution then allowed to run its course to see if predictions were accurate. Our recent and ongoing work provides an alternative approach that is *Bayesian* in nature and allows for DFE updating in real time; it is derived from first principles, it has realistic prospects of practical implementation in real-world settings, and it can be implemented for a small fraction of the cost compared to frequentist-like approaches. We question the utility of DNA sequence data in predicting evolution and join other researchers in pointing to an emphasis on fitness-phenotype evolution as the way forward.

Keywords

Fitness evolution · Phenotype evolution · DNA sequence evolution · Theoretical evolutionary genetics · Mutation · Natural selection · Complexity

1 Introduction

The study of evolutionary processes has been dominated by retrospective analysis based on observations of biological diversity. Methods employed often rely on the analysis of historical data, both observed, e.g., the fossil record, and inferred, e.g., phylogenetic reconstruction. Results of such analysis can tell us about selective pressures, mutational events, the role of genetic drift, migration patterns, and the influence of environmental change, on *past* evolutionary trajectories. However, past evolutionary trajectories are not easily extrapolated to future trajectories, as many of the contingencies shaping past trajectories were impossible to foresee (Nosil et al. 2020; Sarewitza and Pielke 1999).

As such, proposals to develop methods for forecasting evolution have been, and continue to be, a hard sell. Since its inception, evolutionary biology has largely assumed inherent unpredictability to be an inevitable part of its identity as a discipline. Long-term forecasting may indeed be impossible. Extreme short-term forecasting has been studied and practiced by plant and animal breeders for centuries, but breeders have been primarily concerned with predicting phenotypic change over the course of a single generation – a timescale on which artificial selection may be applied but natural selection need not be accounted for. Breeders focus on gene interactions with other genes, phenotypes, and the environment and how these factors might augment desirable characteristics in offspring. The randomness introduced by Mendelian inheritance, however, is unavoidable, making predictions about offspring probabilistic in nature.

The central aim of our contributions to forecasting evolution (Pénisson et al. 2013; Gerrish and Sniegowski 2011, 2012) has been to explore a "goldilocks zone" somewhere between long-term and extreme short-term forecasting: a medium term that (1) leverages the deterministic influence of natural selection for as long as it can provide a useful signal-to-noise ratio and (2) minimizes the probability of major contingencies. The ability to forecast evolution even just a few generations into the future would be invaluable to oncologists, epidemiologists, conservation biologists, and practitioners of infectious disease intervention. Analogies to weather forecasting and financial market forecasting serve to emphasize this point: The economic benefits of accurate short-term forecasting in both cases are enormous.

Benefits of forecasting evolution are not limited to getting ahead of the game in intervention strategies and protocols. Knowledge of evolution's next move could also be used to *steer* evolution in a desirable direction. Ethical issues are obvious and weighty where the evolution of higher organisms is concerned. They are less obvious, and perhaps justifiable, where infectious diseases or cancer is concerned. Previous work has explored the use of mathematical evolutionary forecasting for steering antibody evolution with vaccine protocols toward broadly neutralizing antibodies, in the contexts of HIV and influenza (Gerrish 2016).

1.1 Mathematical Models and Forecasting

It may sound trivial to talk about the relationship between mathematical models and *prediction*. Perhaps this is because many of us think of math and its predictions in terms of physics and engineering: Making predictions about how much weight a bridge can sustain, for example, uses math formulations of physical models. And, thankfully, such *predictions* rarely disappoint.

In an attempt to avoid such preconceptions, the term *forecasting* will be used, because our topic is much more closely related to weather forecasting and financial market forecasting than to physics-based *predictions* used by engineers. The relationship between engineering prediction and physical models is tight. The relationship between evolutionary forecasting and evolutionary models is not. But it is suspected that evolution forecasts, like weather forecasts, can be good enough to be useful.

1.1.1 Hierarchy of Model Agnosticism

When faced with hard forecasting challenges (like the weather), computer scientists sometimes talk about "model-free prediction," usually referring to unsupervised (reinforcement) machine learning. In the language used here, such predictive methods are examples of "uninformed models," because they are initially agnostic, i.e., they initially assume no knowledge of the process to be predicted. Such methods can be likened to maximum likelihood estimation, which is simply Bayesian estimation with a flat (uninformative) prior.

The extreme agnosticism of model-free forecasting implicitly discards centuries of research and hard-won characterizations of nature that have been invaluable to both basic and applied sciences. With no knowledge of Navier–Stokes theory, for example, one would be hard pressed to predict the weather based on model-free forecasting alone (Girault and Raviart 2011).

Most attempts at forecasting employ at least a very basic model of change. The simplest such model is one that assumes no change will take place. In some cases, careful analysis reveals that the process modeled is so random that improvements on the "no change" model cannot be found.

A step above this simplest model is a model that assumes no change in the statistical properties of the process will take place. A few steps above this model is a model that assumes that only one statistical property, namely the expectation, will not change. These models, called *martingales*, form the basis for "detrended" predictive models that can be constructed from a priori knowledge of the process modeled. Such "informed models" are the primary focus of this chapter.

Models based on fitness landscape abstractions have their limitations. Smerlak (2019) notes that populations with different mutation rates or genetic architectures may experience the same fitness landscape in fundamentally different ways, complicating the task of prediction. Additionally, the order in which mutations arise can strongly influence future pathways available on a fitness landscape subsequent evolutionary possibilities, leading to historical contingency and path dependence (Nosil et al. 2020). And environmental fluctuations and gene-by-environment interactions further complicate landscape-based predictive programs (Reeve 2000; Sarewitza and Pielke 1999).

Wortel et al. indicate that public health interventions, such as vaccination and physical distancing, can alter the evolutionary potential of pathogens by modifying population size and selection pressures, thereby influencing the predictability of antigenic evolution. Accounting for such feedback between interventions and evolution is crucial in applying predictive evolutionary biology, as it can inform strategies to manage the emergence of novel variants and guide public health policy (Wortel et al. 2023; Gandon et al. 2016). The transition from retrospective to predictive dynamics in evolution represents a fundamental shift in the field, driven by advances in modeling, computation, and real-time data collection. While retrospective analyses remain invaluable for understanding the history of life, predictive frameworks offer the promise of anticipating and potentially steering evolutionary outcomes (Smerlak 2019; Passagem-Santos et al. 2018; Nosil et al. 2020).

1.2 Darwin's Principle: Fitness Evolution Is a Martingale

1.2.1 Relative Fitness: Evolution's "Potential"

Fitness serves as a unifying concept in evolutionary dynamics, mapping enormous complexity onto a single parameter – a fundamental *potential* – quantifying contribution to future generations. Fitness only makes sense in the context of populations. Darwin's great insight can be summarized quantitatively as noting that fitness is only of consequence in the degree to which it differs from population mean fitness. Such *relative fitness* determines effective fecundity and contribution to future generations. This fact can be understood in the context of traveling wave models, where the distribution of fitness within a population propagates toward ever-increasing *absolute fitness* at constant speed. The wave's constant speed reflects the fact that effective fecundity is determined by relative fitness in finite populations (Luksza and Lässig 2014; Bertram et al. 2017). Students of evolution are well advised to think of absolute fitness as a meaningless quantity most of the time. Lässig et al. (2017) demonstrate the applicability of relative fitness as potential to systems such as influenza A/H3N2, where the interplay between antigenic change and population immunity can be captured by relative fitness dynamics. Despite its challenges, the abstraction of relative fitness as a potential remains a powerful tool for both understanding and predicting evolutionary dynamics. It provides a "potential" for evolution that applies equally to everything from viruses to whales (Wortel et al. 2023; Lässig et al. 2017).

1.2.2 Mean Relative Fitness Is Zero by Definition

The "potential" of Darwinian evolution has an immediately obvious feature that would appear problematic. Mean relative fitness is zero by definition. Letting x denote fitness, then relative fitness is $x - \bar{x}$, where $\bar{x}$ is mean fitness and is a function of continuous time t: $\bar{x} = \bar{x}(t)$. But while the quantity $x - \bar{x}$ is evolution's potential and the driver of evolution, the quantities x and $\bar{x}$ by themselves do not mean much; they are only meaningful in relation to one another. (Several previous authors have pointed out that population mean fitness is, by itself, of little meaning.) Translating these observations to math speak, Darwin's great insight can be stated by saying that **evolution is a martingale**; relative fitness evolution is a process that is inherently and automatically detrended.

1.2.3 Analogies with Physical Potentials

Evolutionary theory, like statistical physics, addresses the behavior of large populations of interacting particles or organisms (Chia and Goldenfeld 2011; Smerlak 2021; Nosil et al. 2020; Gerrish et al. 2013; Smerlak and Youssef 2017; Gerrish and Sniegowski 2012; Lässig et al. 2017; Fisher et al. 2013; Smerlak 2018; Sella and Hirsh 2005; de Vladar and Barton 2011; Matsuda et al. 1992; Barton and Coe 2009; Neher and Shraiman 2011). The notion of a fitness landscape, where genotypes or phenotypes are mapped to a scalar fitness value, closely mirrors the concept of a potential energy surface in physics. This analogy is not merely superficial; it provides a rigorous foundation for modeling evolutionary trajectories as movements within a high-dimensional fitness landscape, akin to particles traversing a potential field. The value of these analogies is twofold. First, they open up possible extrapolation of established techniques from condensed matter and nonequilibrium physics to evolutionary problems. Second, the analogy extends to the interpretation of fitness as a potential that shapes the flow of populations through genotype space. Just as physical systems tend to evolve toward states of lower potential energy, populations under selection are biased toward genotypes of higher fitness. However, unlike many physical systems where parameters remain constant and equilibrium is rapidly achieved, evolutionary systems often operate on longer timescales with fluctuating parameters, making true equilibrium rare or unattainable (Gerrish et al. 2013). This distinction highlights the utility of nonequilibrium statistical mechanics in evolutionary biology, as opposed to the steady-state analyses common in physics (Mustonen and Lässig 2009; Becker and Sibani 2014; Kussell and Vucelja 2014).

Adaptive processes in evolution can be quantitatively described using measures such as the cumulative fitness flux, which captures the total amount of adaptation experienced by a lineage or clade over time. This concept is analogous to the integrated work done by a force in physics, providing a historical record of selective pressures and adaptive events. The fitness flux framework allows for the visualization and quantification of evolutionary trajectories, revealing the underlying structure of adaptation in complex populations (Luksza and Lässig 2014).

While physics analogies provide powerful conceptual and mathematical tools, they also underscore the unique challenges of evolutionary systems. The stochastic emergence of beneficial mutations, the role of genetic drift, tendencies toward criticality, and the influence of fluctuating environments introduce complexities that are less prevalent in physical systems (Figgins and Bedford 2025; Smerlak 2019; Luksza and Lässig 2014; Gerrish et al. 2013).

1.2.4 Lessons from Weather Forecasting

The pursuit of predictability in evolutionary biology has some similarities with the early developments in the field of weather forecasting (Shuman 1989; Namias 1968; Gandon et al. 2016; Cecconi et al. 2012; Peters and Neelin 2006; Richardson 1922). Both disciplines grapple with the challenge of forecasting the future state of complex, high-dimensional systems that are sensitive to initial conditions and subject to stochastic influences (Smerlak 2019). In weather science, the development of mathematical models to simulate atmospheric dynamics, as pioneered by Richardson (1922); Lynch (2016, 2006) and later refined by Lorenz (Shen et al. 2023; Lorenz 1963), has highlighted both the power and the limitations of predictive modeling. Lorenz's work, in particular, shows that even with accurate equations, the intrinsic limits to forecasting arise from the exponential growth of small errors, a phenomenon popularly called the "butterfly effect" (Pielke et al. 2024) that is now formalized through the concept of Lyapunov exponents. This insight has profound implications for evolutionary prediction, where analogous sensitivities can arise.

The world's first numerical weather forecast was computed by Richardson (1922). His 6-hour forecast predicted weather conditions for the city of Munich that were so extreme that they are thought to never have existed on planet Earth

(Richardson 1922; Lynch 2016, 2006). Despite the daunting challenges and a *very* rough start, however, forecasting in meteorology has achieved remarkable success, with the ability to make useful forecasts of weekly weather regimes, monthly means, and even seasonal characteristics improving steadily over the past decades. However, these advances have been contingent on the availability of high-quality data, sophisticated models, and an understanding of the system's inherent unpredictability. The method of analogs, originally applied in weather forecasting, involves identifying past states of the system that closely resemble the current state and using their subsequent evolution as a guide for prediction. While this approach has yielded some success in meteorology, its application in evolutionary biology is complicated by the high dimensionality of genotype space and the context dependence of fitness landscapes (Passagem-Santos et al. 2018; Zhou and McCandlish 2020).

In both fields, the density and quality of observational data play a critical role in determining the accuracy of predictions. For instance, in weather forecasting, the assimilation of real-time atmospheric data has been essential for improving forecast skill. Similarly, in evolutionary studies, real-time population-level data and dense sampling of genotypes are necessary to capture the nuances of fitness landscapes and to enable meaningful forecasts (Zhou and McCandlish 2020; Luksza and Lässig 2014).

A key lesson from weather forecasting is the recognition that not all systems are equally predictable. Some atmospheric phenomena, such as the regular influx of tropical weather patterns, are more amenable to prediction due to strong external forcing, whereas others are subject to a multitude of interacting factors that reduce forecast reliability. This distinction is echoed in evolutionary biology, where certain pathogen dynamics, for example, may be more predictable when driven by potent environmental or demographic factors, while others remain highly contingent on complex genetic and ecological interactions (Gandon et al. 2016; Bisschop et al. 2022).

The forecast horizon, the temporal window over which predictions remain usefully accurate, varies not only with the system but also with the specific variables of interest and the quality of available data. The analogy with weather forecasting also highlights the importance of understanding the limits of predictability. In meteorology, the growth of forecast errors is now a well-characterized phenomenon, and efforts to extend the forecast horizon have focused on improving initial condition estimates, refining models, and quantifying uncertainty (Namias 1968).

Weather forecasting demonstrates that even in the face of intrinsic unpredictability, meaningful and actionable predictions are possible when models are continually updated with new data and when goals for the predictive horizon are modest.

1.2.5 Lessons from Financial Markets

The study of financial market prediction offers a rich source of analogies and methodological insights for evolutionary forecasting (Levin and Lo 2021; Diep and Desgranges 2021; Amir et al. 2021; Robson and Orr 2021; Zhou 2016; Autzen 2024; Koduri and Lo 2021). Both fields grapple with the challenge of predicting the future behavior of complex, stochastic systems, where the underlying dynamics are shaped by a combination of deterministic rules and random fluctuations. In financial markets, the efficient market hypothesis (EMH) posits that all available information is instantaneously reflected in asset prices, rendering past price movements largely uninformative for future predictions. This principle has profound implications for modeling, as it suggests that only new, unforeseen information can drive price changes and that the best predictor of tomorrow's price is today's price. The EMH thus forms a martingale as does the relative fitness process in evolution; in addition, it has the Markov memoryless property, where the system's future state depends solely on its present state, not its history. This perspective resonates with certain approaches in evolutionary biology, particularly those that emphasize the primacy of current population states and fitness landscapes in shaping future evolutionary trajectories. In both domains, the challenge lies in extracting predictive signals from noisy, high-dimensional data, while accounting for the fact that the system is constantly being perturbed by novel inputs, be they mutations in evolution or news in financial markets.

The analogy extends further: Just as financial models often rely on stochastic processes such as geometric Brownian motion to capture the random walk nature of asset prices, evolutionary models frequently employ stochastic differential equations to describe allele frequency changes under the influence of mutation, selection, and drift. However, the limitations of financial market prediction also serve as a cautionary tale for evolutionary forecasting. The EMH implies that, in an idealized market, systematic prediction is fundamentally constrained by the rapid incorporation of information, leading to a situation where only unpredictable, exogenous shocks can move prices (Dmouj 2006). Similarly, in evolution, the unpredictability of future mutational inputs and environmental changes can limit the accuracy of forecasts, especially over longer timescales. Nosil et al. (2020) indicate that rare and difficult-to-predict environmental events can have outsized effects on evolutionary outcomes, mirroring the impact of black swan events in financial markets (Box 3).

Despite these challenges, both fields have developed sophisticated statistical and computational tools to improve short-term predictive power. In finance, Bayesian estimation and real-time updating of model parameters are standard techniques for assimilating new information and refining forecasts (Dmouj 2006). Evolutionary biology is increasingly adopting similar approaches, using Bayesian inference to estimate the distribution of fitness effects (DFEs) and to update predictions as new genetic data become available. Wortel et al. (2023) state that predicting evolutionary outcomes often involves estimating the probability of specific events, such as the fixation of beneficial mutations or changes in population size, which parallels the probabilistic forecasting of asset returns or market crashes.

A key insight from financial modeling is the importance of understanding the structure and flow of information within the system. In markets, transparency and the rapid dissemination of information are essential for maintaining efficiency and limiting arbitrage opportunities (Dmouj 2006). In evolutionary systems, genetic linkage and the structure of genealogical trees play analogous roles in preserving and transmitting information about the DFE and the fitness landscape. According to Luksza and Lässig (2014), the inference of genealogical trees and the estimation of clade frequencies are central to predicting evolutionary dynamics, much as the analysis of order books and transaction histories informs market predictions. Furthermore, the analogy highlights the value of focusing on short-term dynamics, where predictive models are most likely to succeed.

In finance, short-term forecasts can exploit transient inefficiencies or patterns before they are arbitraged away, while long-term predictions are hampered by the accumulation of unpredictable shocks (Dmouj 2006). In evolution, models that prioritize short-term fitness trajectories and real-time estimation of mutational inputs are better positioned to capture the immediate consequences of selection and drift, even if long-term outcomes remain elusive. The work of Luksza and Lässig (2014) demonstrates that extending prediction schemes from population frequencies to absolute growth rates and incidence rates can enhance the accuracy of short-term forecasts, provided that the relevant selective effects are well characterized.

The lessons from financial market prediction thus underscore both the promise and the limitations of evolutionary forecasting. They suggest that while perfect prediction may be unattainable due to the inherent stochasticity and information constraints of complex systems, meaningful progress can be made by leveraging real-time data, focusing on short-term dynamics, and rigorously quantifying uncertainty. The cross-fertilization of ideas between these fields continues to inspire new approaches to the fundamental problem of predicting the future in systems governed by both chance and necessity (Dmouj 2006; Nosil et al. 2020; Wortel et al. 2023; Luksza and Lässig 2014).

2 Theory

2.1 Models of Fitness Evolution

Standard models of fitness evolution are typically grounded in the dynamics of genotype frequencies within a population, where the change in log-frequency of a genotype is determined by the difference between its fitness and the mean fitness of the population. Mutation is often modeled as a diffusion process, with a strong bias toward deleterious effects, reflecting the empirical observation that most mutations are harmful rather than beneficial. The distribution of fitness effects (DFEs) of new mutations is a key determinant of evolutionary trajectories. The fact that the DFE can often be assumed to be stationary over short timescales enables the use of time series analyses to make predictions about future evolutionary states. However, periods of non-stationarity, where the DFE shifts due to environmental changes or internal population dynamics, can signal impending evolutionary transitions, such as the onset of metastasis in cancer or host shifts in pathogens (Gerrish and Hengartner 2017). The ability to detect statistical rumblings in the DFE that commonly precede such “black swan events” (Box 3) could inform conservation measures and public health decisions.

Gene regulatory networks can evolve new connections while maintaining ancestral functions, illustrating how evolutionary complexity is constrained by both the need to preserve essential traits and the opportunity to explore new adaptive peaks. These constraints generate correlations in phenotypic evolution that can be exploited for inferring fitness landscapes and for predictive analyses. Mechanistic fitness models, which incorporate both positive and negative fitness components, are feasible when a few key phenotypes explain most of the fitness variation in the population. Such models enhance predictability by reducing the effective dimensionality of the evolutionary process (Lässig et al. 2017). In the context of quantitative genetics, forecasting approaches often rely on frameworks such as the breeder's equation or the Robertson–Price identity, which relate the response to selection to phenotypic and genetic variances and covariances. These models assume that the genetic architecture of traits and the covariance between traits and fitness are sufficiently stable to allow for meaningful predictions of evolutionary responses (Wortel et al. 2023).

The temporal dynamics of fitness in evolving populations are shaped by a complex interplay of mutational input, selection, genetic drift, and linkage, all of which contribute to the patterns and predictability of evolutionary change. Recent quantitative models have increasingly focused on capturing these dynamics in real time, moving beyond static or retrospective analyses to frameworks that can forecast short-term evolutionary trajectories based on current population data. In high mutation regimes, linkage can generate complex, sometimes counterintuitive, evolutionary dynamics that deviate from the expectations of models assuming free recombination. Mechanistic fitness models, as outlined by Lässig et al. (2017), are particularly effective when a few key phenotypes explain most of the fitness variation within a population. These models typically incorporate both positive and negative fitness components, which together constrain the evolutionary complexity and enhance the predictability of fitness trajectories. The temporal evolution of fitness in such models is not only merely a function of the accumulation of beneficial mutations but also reflects the balance between adaptation and the costs associated with resource allocation, as seen in microbial systems where compound production diverts resources from growth.

Polygenic forecasting methods, which leverage quantitative genetics frameworks such as the breeder's equation and the Robertson–Price identity, offer additional tools for predicting the response to selection based on phenotypic and genetic variances and covariances (Wortel et al. 2023). These approaches are particularly relevant when adaptation involves multiple traits or when the genetic architecture of fitness is highly polygenic. Gallegos et al. (2025) demonstrate the application of quantitative genetic models to simulate the evolution of environmental tolerance via adaptive changes in plastic traits, illustrating how temporal series data can inform predictions of fitness evolution in complex environments. The integration of deterministic and stochastic models further enriches our understanding of temporal fitness dynamics. Martin and Roques (2016) discuss the use of deterministic approximations in non-epistatic models to retrieve known properties and generate new insights into the temporal progression of fitness. However, the extension of these models to account for epistatic interactions and context-dependent mutation effects remains a significant challenge.

In the context of infectious disease evolution, the temporal dynamics of fitness are particularly salient. Gandon et al. (2016) emphasize that forecasting the dynamics of pathogens such as influenza requires models that incorporate both epidemiological and evolutionary dimensions, as the fitness landscape of the pathogen is continually reshaped by host immunity and intervention strategies. The collective insights from these studies indicate that while significant progress has been made in modeling and forecasting the temporal dynamics of fitness, challenges persist in scaling these approaches to encompass multiple phenotypes, complex genetic architectures, and the full spectrum of evolutionary forces. Nonetheless, the ongoing refinement of quantitative models, informed by real-time data and advanced computational methods, holds promise for improving the predictability of evolutionary outcomes (Nosil et al. 2020; Lässig et al. 2017).

2.2 Estimating Mutational Input

Mutational input is characterized by two key components: the mutation rate and the distribution of fitness effects (DFEs) of new mutations. Estimation of these components is complicated by the fact that they are not independent of each other. On the surface it seems plausible, for example, that a high rate of production of mutations of small effect might create evolutionary dynamics similar to – and perhaps statistically indistinguishable from – a low rate of production of mutations of large effects. Indeed, for fully nonparametric estimation, it has been shown (Gerrish and Hengartner 2017) that it is impossible to estimate both the mutation rate and the DFE – a sort of evolutionary uncertainty principle.

2.2.1 Mutation Rate

Mutation rate estimation is a central component in characterizing the mutational input to evolutionary processes, as it directly determines the frequency and spectrum of new genetic variants entering a population. Accurate estimation of mutation rates is essential for understanding the distribution of fitness effects (DFEs) and for constructing predictive models of evolutionary dynamics. The DFE itself represents the statistical distribution of the fitness consequences of new mutations, and its accurate characterization depends critically on reliable mutation rate estimates (Gerrish and Hengartner 2017). The estimation of mutation rates is complicated by the need to disentangle the effects of selection from the raw mutational input. Natural selection acts as a strong filter, biasing the observed spectrum of mutations away from the true underlying distribution.

2.2.2 Distribution of Fitness Effects (DFEs)

Mutations that alter fitness serve as the fundamental input for natural selection, yet the general properties of these mutations, such as the proportion that is beneficial versus deleterious and the magnitude of their effects, remain incompletely characterized. The DFE represents the spectrum of possible fitness consequences arising from genetic changes, and its shape directly influences the rate and trajectory of adaptation, the likelihood of parallel evolution, and the maintenance of genetic variation within populations. Recent theoretical and empirical efforts, by ourselves and others, have focused on quantifying the DFE, particularly for beneficial mutations (Aguilar 2025; Gerrish and Hengartner 2017), due to its implications for forecasting evolutionary outcomes. However, while the overall distribution can be estimated, connecting specific mutations to their precise fitness effects and their probabilities of occurrence remains a significant challenge. This limitation arises because the DFE is not a static property; it is shaped by the underlying genotype-to-phenotype map, the genetic background, and the environment in which selection occurs. Lind et al. (2019) note that models incorporating increasing levels of information about genetic architecture can improve explanatory power, but minimal information is still required for reliable predictions. The DFE is not only a property of individual mutations but also reflects the structure of the fitness landscape. For example, the anisotropy in the spectral distribution of mutational effects demonstrates that the directionality and magnitude of fitness changes can vary depending on the position in phenotype space and proximity to fitness optima. This anisotropy implies that the DFE may shift as populations approach or move away from adaptive peaks, with the distribution of effects changing in both mean and variance (Martin 2014).

Empirical studies in microbial populations have provided insights into the DFE by measuring the fitness effects of fixed beneficial mutations (Rozen et al. 2002). The shape of the DFE's beneficial tail, whether it is exponential, heavy-tailed, or otherwise, has important consequences for the predictability of adaptation and the potential for parallel evolution (Heilbron et al. 2014; Lind et al. 2019). Furthermore, the fixation of weakly and highly deleterious mutations, especially in populations experiencing recurrent bottlenecks or elevated mutation rates, can lead to punctuated declines in fitness, as observed by Heilbron et al. (2014). Mathematical models, such as those developed by Gerrish and colleagues (Gerrish and Hengartner 2017; Gerrish and Sniegowski 2012), employ recurrence relations to map the cumulants of fitness distributions over time to the genomic mutation rate and the moments of the DFE.

Despite advances in quantifying the DFE and integrating it into predictive models, significant challenges remain. The extension of these approaches to systems with multiple phenotypes or to sequence-level evolution is complicated by the increased dimensionality and the complexity of genotype–phenotype–fitness relationships (Lind et al. 2019; Martin 2014). Moreover, the DFE is context-dependent, varying with environmental conditions, genetic background, and the evolutionary history of the population. Yet its accurate characterization is essential for developing robust mathematical frameworks capable of predicting evolutionary change in real time (Lind et al. 2019; Gerrish et al. 2013; Martin 2014).

Box 1 Fitness–Phenotype Evolution
Let random variable X denote fitness and let random vector $\mathbf{Y}$ consist of numerical values of measurable phenotypes thought to affect fitness. In previous work, we have shown that X and $\mathbf{Y}$ evolve as

$$\dot{\mathcal{C}} = \underbrace{\mathcal{C}' - \mathcal{C}'(0)}_{selection} + \underbrace{U(\mathcal{M} - 1)}_{mutation} + \underbrace{(1 - e^{\mathcal{C}^* - \mathcal{C}})/2N_e}_{genetic\ drift} \tag{1}$$

where:

- $\mathcal{C} = \mathcal{C}(\theta, \mathbf{\Phi}, t) = \log \mathbb{E}[e^{\theta X + \mathbf{\Phi}^{\mathbf{T}}\mathbf{Y}}]$, the CGF of fitness and phenotypes.
- $\mathcal{M} = \mathcal{M}(\theta, \mathbf{\Phi}, t) = \mathbb{E}[e^{\theta \Delta X + \mathbf{\Phi}^{\mathbf{T}}\Delta\mathbf{Y}}]$, the MGF of the DFE.
- $\mathcal{C}^* = \mathcal{C}(2\theta, 2\mathbf{\Phi}, t)$; $\dot{\mathcal{C}} = \partial_t \mathcal{C}$; $\mathcal{C}' = \partial_\theta \mathcal{C}$; $\mathcal{C}'(0) = \partial_\theta \mathcal{C}|_{\theta=0}$
- θ and $\mathbf{\Phi}$ are dummy variable and vector.
- U is genomic mutation rate, and N_e is effective population size.

The non-locality of $\mathcal{C}^*$ complicates analytical solution of Eq. (1), but if the genetic drift term is ignored, i.e., under the reasonable assumption that N_e is large, Eq. (1) becomes a variant of the transport equation with solution

$$\mathcal{C}_t(\theta) = \mathcal{C}_0(\theta + t) - \mathcal{C}_0(\theta) + U \int_0^t \mathcal{M}_t(\theta + \gamma) - \mathcal{M}_t(\gamma) d\gamma \tag{2}$$

where

- $\mathcal{C}_t(\theta) = \mathcal{C}(\theta, \mathbf{\Phi}, t)$, and $\mathcal{M}_t(\theta) = \mathcal{M}(\theta, \mathbf{\Phi}, t)$.

2.3 Fitness Measurement

To make the mutational input estimates required for evolutionary forecasting, as outlined above, we need a way to measure fitness of individuals in the evolving population. Serious endeavors to measure fitness, however, quickly lead to deep questions about what fitness is and to a realization that quantities we can measure are only proxies of absolute fitness. These include maximum exponential growth rate, lag phase (in bacteria), total carrying capacity, competitive fitness as measured by head-to-head competition with a common ancestor or other reference strain, etc. Whatever the measure, it must be treated as a proxy of absolute fitness

(Perfeito et al. 2007, 2011; Trindade et al. 2010; Passagem-Santos et al. 2018); specifically, the mean measure must be subtracted from individual measures to calculate relative fitness, evolution's *potential*.

2.3.1 Experimental Evolution and Real-Time Sampling

Experimental evolution, when combined with real-time sampling, has become a cornerstone for quantifying evolutionary dynamics and testing the central tenets of evolution in a controlled context. It is in this context that the predictive power of mathematical/computational models of evolution can be assessed in a truly closed-book way. These approaches are particularly promising for capturing the temporal dynamics of the distribution of fitness effects (DFEs). High-throughput competitive fitness assays facilitate the measurement of fitness across large numbers of clones, enabling the estimation of the DFE in real time.

Methods outlined in Box 2 provide a direct way to estimate the DFE in transformed spaces, especially Laplace, Fourier, and wavelet spaces, from which valuable information can be extracted, such as evolutionarily important tail behavior. The absence of a plot graphing the DFE directly in delta-fitness space, however, can limit intuitive interpretation. This limitation can be remedied by applying back-transformation methods such as the Fourier transform, saddle point, and other back-transformation methods.

New experimental methods to track frequencies of individual genotypes in vitro or in vivo offer promising new ways to measure fitness in situ and to measure its evolution in real time. These approaches often involve the use of barcoded genotypes and/or deep sequencing; they enable the direct observation of clade dynamics and the validation of fitness models by comparing predicted and actual evolutionary outcomes. Despite these advances, significant challenges remain in extending experimental evolution and real-time sampling to more complex scenarios, such as the evolution of multiple phenotypes or sequence-based traits. The genetic architecture of many traits, especially in eukaryotes, is highly polygenic, complicating the mapping between genotype and phenotype and increasing the dimensionality of the fitness landscape. Wortel et al. argue that detailed knowledge of the genotype–phenotype map becomes increasingly important when attempting to control or forecast evolutionary responses in such systems. Furthermore, the presence of epistasis and the structure of the fitness landscape can constrain the number of accessible evolutionary pathways, thereby influencing the predictability of adaptation. For example, sign epistasis can reduce the number of mutational routes to higher fitness, increasing the repeatability of evolutionary trajectories and enhancing the prospects for forecasting (Wortel et al. 2023). However, the interplay among epistasis, linkage, and environmental fluctuations introduces additional layers of complexity that must be accounted for in both experimental design and data analysis. In summary, experimental evolution combined with real-time sampling provides a powerful platform for measuring and modeling evolutionary dynamics and assessing predictive frameworks (Gerrish and Hengartner 2017; Wortel et al. 2023; Luksza and Lässig 2014).

Box 2 A Forecasting Procedure

1. **Data collection**. At timepoint k, n individuals are sampled from an evolving population. Fitness x and a number of phenotypes $\mathbf{y}$ are measured that are thought to be related to fitness. This yields dataset $\mathcal{D}_k = \{\{x_1, \mathbf{y}_1\}, \{x_2, \mathbf{y}_2\}, ..., \{x_n, \mathbf{y}_n\}\}_k$.
2. **Estimating the DFE**. The MGF of the DFE, $\mathcal{M}$, is obtained by rearranging Eq. (1) (the "derivative method") or Eq. (2) (the "integral method"). Strictly speaking, only the quantity $U(\mathcal{M} - 1)$ can be estimated (nonparametrically), as follows:

$$U(\tilde{\mathcal{M}} - 1) = F_D(\tilde{\mathcal{C}}) \quad \text{or} \quad U(\tilde{\mathcal{M}} - 1) = F_I(\tilde{\mathcal{C}})$$

   where functions F_D and F_I are derived from the derivative and integral method, respectively. The arguments are empirical CGFs (or ECGFs), defined as

$$\tilde{\mathcal{C}} = \log \frac{1}{n} \sum_{i=1}^{n} e^{\theta x_i + \mathbf{\Phi^T y_i}}$$

3. **Updating the DFE**. The posterior MGF of the DFE can be computed by bootstrapping on the dataset, $\mathcal{D}_k$:

$$g(\mathcal{M}_k(\theta)) = f(\mathcal{M}_k(\theta)|\mathcal{D}_k^b) \propto f(\mathcal{D}_k^b|\mathcal{M}_k(\theta))g(\mathcal{M}_{k-1}(\theta)) \ \forall\, \theta$$

   where the prior MGF is simply the posterior from the previous sample.
4. **Predicting future state**. The random vector of fitness and phenotypes at future timepoint $k + 1$ is given by

$$\mathcal{D}_{k+1} \sim f(\mathcal{D}|\mathcal{M}_k(\theta))g(\mathcal{M}_k(\theta))$$

5. **Variants of this procedure**. While all past states are implicitly accounted for by this procedure, the weights of past states can be increased by defining the prior to be an average of, say, the two prior states. Alternatively, or additionally, temporal trends in quantiles of the priors may be fitted to a time series model and used in the forecasting procedure.

2.3.2 Measurement Error

The accuracy of evolutionary predictions is fundamentally constrained by the quality and reliability of the data used to parameterize and validate models. In evolutionary experiments, fitness and genotype or phenotype measurements are often subject to various sources of noise, including technical limitations of measurement devices, biological variability, and sampling error. These sources of error can obscure true evolutionary signals and complicate the inference of underlying processes. A key advantage of formulating evolutionary models in terms of cumulants is their capacity to explicitly account for additive measurement errors. When fitness or phenotype measurements are contaminated by noise with a constant distribution, the cumulants of the observed data become the sum of the true cumulants and those of the error distribution. This additive property allows for a more transparent separation of biological signal from measurement noise, facilitating more robust statistical inference and model comparison. Because of this additive property, the protocol we outline in Box 2 turns out to be inherently robust to measurement noise when the math outlined in Box 1 is used for both estimating the DFE and forecasting evolution; the noise terms simply cancel out.

The challenge of measurement error is further compounded when comparing the fit of different evolutionary models. Metrics such as r^2 quantify the proportion of variance in the observed data explained by a model, but they do not capture the model's ability to reproduce the stochasticity inherent in evolutionary processes. Cross-correlation measures, on the other hand, assess the correspondence between observed and predicted trajectories, including their stochastic fluctuations. These complementary statistics are essential for evaluating model adequacy, especially when the summary statistics used for model selection are the same as those reported for evolutionary trajectories, such as fitness measures and the number of acquired mutations. According to Passagem-Santos et al. (2018), the use of both r^2 and cross-correlation provides an improved assessment of model performance, highlighting the importance of capturing both deterministic trends and stochastic variability. In experimental studies, the variance attributable to measurement error can differ across treatments or experimental conditions. For example, when the variance due to measurement error is consistent among treatments, the relative contribution of selection to total variance may appear reduced after events such as population bottlenecks. This observation underscores the necessity of carefully partitioning sources of variance to avoid misattributing changes in evolutionary dynamics to biological processes when they may instead reflect shifts in measurement error (Bisschop et al. 2022).

Wortel et al. (2023) state that the context-dependent importance of different evolutionary parameters, such as fitness or mutational input, further complicates the interpretation of noisy data, as the relevance of specific measurements may shift depending on the evolutionary scenario under investigation. The process of estimating selection coefficients and distributions of fitness effects (DFEs) is also sensitive to measurement error. In some studies, measurement error is neglected to focus on qualitative agreement between theory and data, but this simplification may limit the generalizability of findings (Martin 2014). Martin and Roques (2016) suggest that the accuracy of deterministic approximations to stochastic evolutionary models is often evaluated using simulations, which can help quantify the error introduced by such approximations. However, the translation of these results to empirical data is contingent on the reliability of the underlying measurements. Cross-validation techniques, such as dividing observed time series into training and test sets, are commonly employed to assess the predictive performance of evolutionary models. Performance metrics like the correlation coefficient (ρ) and normalized root mean square error (RMSE) are used to quantify the agreement between observed and predicted values. These metrics provide complementary perspectives on model accuracy, with correlation capturing the strength of association and RMSE reflecting the magnitude of prediction errors.

Wu et al. (2024) emphasize that the design of unified and reliable frameworks for forecasting in complex systems remains an open challenge, in part due to the difficulties associated with measurement error and the high dimensionality of biological data. The use of direct measurement techniques, such as site-directed mutagenesis followed by competitive fitness assays, can reduce the detection limit for selection coefficients, thereby improving the resolution of DFE estimates (Martin 2014). Nevertheless, even with precise measurements, the neglect of measurement error in analysis can bias conclusions about the qualitative agreement between theoretical predictions and empirical data. The iterative process of model fitting, validation, and refinement thus requires careful attention to the statistical properties of measurement error and its impact on parameter estimation. The development of robust mathematical frameworks necessitates explicit strategies for quantifying and correcting for measurement noise, the use of multiple complementary performance metrics, and the careful interpretation of variance components in experimental data (Wortel et al. 2023; Martin and Roques 2016; Martin 2014; Passagem-Santos et al. 2018; Wu et al. 2024; Bisschop et al. 2022). These considerations are essential for advancing the predictive power of evolutionary models and for ensuring that forecasts are both accurate and biologically meaningful.

Box 3 Agnostic Threat-Level Monitoring for Black Swan Events

1. **Nonparametric detection of structural changes in the DFE**. A major change in $\mathcal{M}(\theta)$ could be a harbinger of the kind of major evolutionary change that can underlie worrisome events like zoonoses. The methods outlined in Box 2 can be used to monitor for such harbingers. For example, threat levels might be raised when $\mathcal{M}_k(\theta)$ lies outside of the 95% confidence interval of the null ensemble of functions on θ constructed by bootstrapping on $\mathcal{M}^*(\theta)$, averaged over different values of θ, where $\mathcal{M}^*$ is some function of previous $\mathcal{M}$'s; $\mathcal{M}^* = \mathcal{M}_{k-1}$ and $\mathcal{M}^* = \mathcal{M}_1$ are two simple examples. (Alternatively, Mahalanobis distance may be used, but we have found that the attendant covariance matrix often has an undefined inverse.)
2. **A numerical example**. Simulations of evolving populations were performed. At the temporal midpoint, a black swan event was simulated by changing the DFE from a Gaussian to a uniform distribution, taking care to make sure that the probability mass above zero remained unchanged. (Evolutionary dynamics are extremely sensitive to this quantity.) Pairwise comparisons are plotted in Fig. 1.

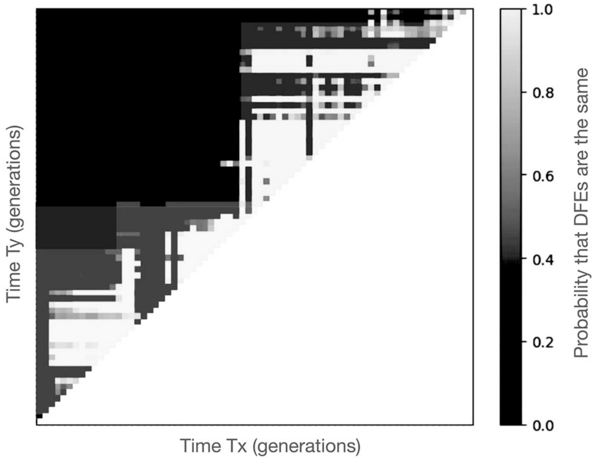

Fig. 1 Monitoring for structural changes in the DFE. Horizontal and vertical axes show indices i and j, respectively, for nonparametric pairwise comparison between H0: $\mathcal{M}^* = \mathcal{M}_i(\theta)$ and H1: $\mathcal{M}_j(\theta)$ $(j > i)$; black indicates they are statistically distinguishable ($p < 0.05$). Here, our methods successfully detect a simulated black swan event introduced at the halfway point

2.4 Genetic Linkage

The dynamics of fitness distributions in evolving populations are shaped by the interplay between genetic linkage and recombination. Genetic linkage, by constraining the independent assortment of alleles, preserves correlations among loci and thus maintains information about the distribution of fitness effects (DFEs) across the genome. This preservation is especially relevant in asexual or low-recombination populations, where linkage disequilibrium can lead to the collective behavior of mutations, resulting in complex evolutionary trajectories that deviate from expectations of models assuming free recombination (Couce et al. 2024; Bertram et al. 2017; Merleau et al. 2021). Under linkage, the fate of a beneficial or deleterious mutation is determined solely not only by its own effect but also by the genetic background on which it arises, leading to phenomena such as Muller's ratchet (Colato and Fontanari 2001; Fontanari et al. 2003; Haigh 1978; Colato et al. 2005), lineage contamination (Pénisson et al. 2013, 2017; Bräutigam and Smerlak 2022), and clonal interference and hitchhiking (Sniegowski et al. 1997; Pénisson et al. 2017; Miralles et al. 1999; de Visser et al. 1999; Rozen et al. 2002; Sniegowski and Gerrish 2010; Gerrish and Lenski 1998; Gerrish 2001; Pénisson et al. 2013), which can skew the distribution of fitness within the population (Couce et al. 2024).

Such contingencies arising from linkage are examples of epistasis, where the effect of a mutation depends on the genetic context. Bertram et al. (2017) indicate that macroscopic epistasis can alter long-term evolutionary dynamics by modifying the shape and variance of the fitness distribution, potentially leading to nonlinear responses to selection and unpredictable evolutionary outcomes. In tightly linked genomes, such as those of many microbes, the accumulation of epistatic interactions can result in rugged fitness landscapes, where populations may become trapped on local fitness peaks, thereby affecting both the mean and variance of fitness over time (Couce et al. 2024). This effect is less pronounced in populations with high recombination rates, where genetic shuffling can break up unfavorable combinations and facilitate the exploration of new adaptive peaks. Recombination acts as a counterbalance to linkage by generating novel allele combinations, thereby increasing genetic diversity and enabling more efficient selection on individual loci.

Wortel et al. state that explicit population genetic models are required to account for the stochastic forces, such as drift, migration, and mutation, that interact with recombination to shape the evolutionary process. The introduction of recombination can reduce the impact of clonal interference, allowing beneficial mutations to combine and spread more rapidly, which in turn can accelerate adaptation and alter the trajectory of the fitness distribution (Wortel et al. 2023). However, the benefits of recombination are context-dependent; in some cases, recombination can break up co-adapted gene complexes, leading to a temporary reduction in mean fitness (Bertram et al. 2017).

Mean fitness flux, as described by Luksza and Lässig (2014), serves as a quantitative measure of the correlation between predicted and observed changes in clade frequencies, offering a metric for assessing the quality of evolutionary predictions in the presence of linkage. Experimental evolution studies have demonstrated that the structure of the fitness distribution is not static but evolves in response to the interplay among mutation, selection, and linkage. For instance, Couce et al. (2024) discuss how initial mutations can direct alternative pathways of protein evolution.

2.4.1 Informative Higher Cumulants

Genetic linkage, while often detrimental to efficient adaptation, can increase the predictability of evolution. This is because linkage can have the effect of maintaining the statistical significance of higher cumulants in the fitness distribution. Higher cumulants, such as skewness and kurtosis, and their dynamics, carry information about the shape and evolution of the distribution of fitness effects (DFEs) that is missing when only the mean and variance of the fitness distribution are significant. In populations where genetic linkage is strong, the nonrandom association of alleles at different loci can preserve statistical information stored in the higher cumulants.

3 Phenotypic Evolution

3.1 Mapping Phenotype to Fitness

Mapping fitness to phenotype is a central challenge in evolutionary biology, as it requires translating genetic and phenotypic variation into quantitative predictions of evolutionary trajectories. The genotype–phenotype–fitness map is inherently complex, shaped by pleiotropy, epistasis, and environmental interactions, which together define the structure of the fitness landscape (Martin 2014; Wortel et al. 2023). Recent advances in high-throughput experimental techniques, such as deep mutational scanning, have enabled the empirical mapping of genotype–phenotype relationships for specific molecular systems. For example, fitness landscapes for protein folding or binding often exhibit a "mesa" structure, with plateaus at high and low folding probabilities and a sharp transition region where fitness changes rapidly with small changes in phenotype. These landscapes can sometimes be parameterized with relatively few variables, allowing for predictive modeling when sufficient training data are available. However, the challenge remains to generalize these mappings to more complex, multidimensional phenotypes and to whole-organism fitness, where the number of relevant variables and interactions increases dramatically (Lässig et al. 2017).

3.2 Phenotype Dynamics

Temporal dynamics of phenotypic traits are central to understanding and predicting evolutionary processes, particularly when considering the interplay among mutational inputs, selection, and genetic linkage. The evolution of phenotypic traits is inherently a dynamic process, shaped by the continuous introduction of mutations and their subsequent effects on fitness within a population. The mathematical treatment of these dynamics often involves tracking the joint distribution of fitness and associated phenotypes over time, as formalized by the use of log-characteristic functions such as $C(\Theta, \Lambda) = \log \mathbb{E}(e^{i\Theta X + i\Lambda Y})$, where X denotes fitness and Y a fitness-related phenotype (Gerrish and Hengartner 2017). This approach enables the quantification of phenotypic shifts in response to both stochastic mutational events and deterministic selection pressures.

The temporal evolution of phenotypic means and variances is a function not only of the underlying genetic architecture but also of the environmental context and the structure of the fitness landscape. For instance, in static environments, the supply of beneficial mutations is finite, leading to a gradual depletion as adaptation proceeds. This results in a characteristic trajectory where mean fitness initially increases but may plateau or even decline as the pool of available advantageous mutations diminishes, with corresponding changes in the variance of fitness and phenotypic traits (Gerrish et al. 2013). Such dynamics underscore the importance of considering both the short-term and long-term consequences of adaptation, as the predictability of phenotypic evolution is often highest in the early stages when the distribution of fitness effects (DFEs) is well characterized, and the population is far from any fitness peak (Couce et al. 2024).

3.3 Multivariate Phenotypic Evolution

3.3.1 Correlated Trait Evolution

Correlated trait evolution arises when multiple phenotypic traits do not evolve independently but instead exhibit statistical associations due to shared genetic, developmental, or selective factors. In multivariate evolutionary models, such as extensions of Fisher's geometric model (FGM), the joint dynamics of several traits are captured by considering the covariance structure among them. This approach allows for a more nuanced understanding of how selection, mutation, and genetic drift shape the distribution of phenotypes in a population (Martin 2014).

The generalized FGM framework, as discussed by Martin, extends classical results by allowing the phenotypic variance–covariance matrix to vary over time, rather than assuming it remains constant. This is particularly relevant when the mutation rate or the strength of selection changes or when the population is far from equilibrium (Gerrish et al. 2007). The model provides explicit thresholds for when the approximation of independence between mean and variance holds and when correlations between traits become significant for predicting evolutionary trajectories (Martin and Roques 2016; Martin 2014). In this context, the evolution of the mean phenotype vector is influenced not only by the direct effects of selection on each trait but also by the structure of mutational covariances, which can channel evolutionary change along certain directions in trait space.

The presence of genetic linkage complicates the interpretation of observed correlated trait evolution, which can derive from physical genetic linkage or functional interdependence (epistasis). Empirical and theoretical studies have shown that negative epistasis among beneficial mutations, where the combined effect of two mutations is less than the sum of their individual effects, can lead to diminishing returns as populations approach local fitness optima. This phenomenon is often observed in multivariate trait evolution, where the fitness landscape is curved and the effects of mutations depend on the current position in phenotype space (Martin 2014). The local approximation in such models assumes that the parental phenotype is not too distant from the optimum, allowing for linearization of the fitness surface and tractable predictions of trait evolution.

Recent advances in modeling correlated trait evolution have also incorporated the concept of fitness flux, which quantifies the cumulative amount of adaptation in a population over time. By extending fitness flux theory to account for clade-dependent and multivariate fluxes, researchers can map the adaptive history of lineages and compare the predictive power of different fitness models (Luksza and Lässig 2014). This approach is particularly useful for understanding the adaptive landscape of rapidly evolving organisms, such as influenza, where multiple traits (e.g., antigenic properties and replication rate) evolve in concert. Experimental evolution studies have begun to test these theoretical predictions by manipulating the covariance structure among traits and observing the resulting evolutionary dynamics in controlled laboratory settings. Such experiments provide critical data for refining multivariate models and for assessing the extent to which correlated trait evolution can be predicted from first principles (Wortel et al. 2023). The interplay among mutation, selection, and genetic architecture thus remains a central focus in the study of multivariate phenotypic evolution, with ongoing efforts to extend these models to more complex scenarios involving sequence evolution and higher order epistatic interactions (Zhou and McCandlish 2020; Smerlak 2021).

4 The Longshot: Forecasting Sequence Evolution

The astronomical dimensionality of DNA sequence space makes talk of forecasting sequence evolution seem laughably unrealistic. However, genetic sequences of a population of organisms of the same species occupy only a tiny fraction of this space. And within that population, the sequences that are ultimately of evolutionary consequence occupy only a tiny fraction of that tiny fraction. A small number of sequences that are left feed the right tail of the DFE. The output of evolutionary processes is constrained by the mutational input: Novel genetic material cannot arise without corresponding mutations that occur prior to, and independently of, natural selection. This places the DFE at the heart of evolutionary forecasting, as it determines the spectrum of possible fitness changes accessible to a population. However, estimating the DFE remains methodologically challenging, especially in natural populations and in real time as our Bayesian framework requires.

Booker (2020) suggests that while precise estimates of the DFE can be obtained from the unfolded site frequency spectrum (uSFS), biases may arise when beneficial mutations are both rare and strongly selected, underscoring the need for robust statistical frameworks. Genetic linkage further complicates sequence evolution by preserving correlations between loci, thereby affecting the fate of new mutations and the repeatability of evolutionary paths (Luksza and Lässig 2014). The work of Lässig et al. emphasizes that redundancies in regulatory sequences and metabolic pathways expand the mutational target size, making genome evolution at the sequence level less repeatable and more contingent on stochastic events. This is contrasted with phenotypic evolution, where certain adaptive outcomes, such as immune receptor–antigen affinities, can be more predictable due to the convergence on specific sequence motifs. Lässig et al. (2017) indicate that while selection reduces the number of viable evolutionary trajectories, the vastness of sequence space ensures that sequence evolution remains largely unrepeatable, even under strong selection.

Despite these challenges, empirical studies leveraging high-throughput sequencing and experimental evolution have begun to map the fitness effects of thousands of mutations in real time, providing unprecedented resolution of the short-term dynamics of sequence evolution. For instance, Couce et al. (2024) used bulk competition assays to estimate selection coefficients for a large fraction of possible gene disruptions, revealing the distribution of fitness effects across the genome.

The unpredictability of sequence evolution is not absolute. While the sheer number of possible mutational paths and the influence of stochastic events limit long-term forecasts, focusing on short-term fitness trajectories and incorporating information about genetic linkage and the DFE can improve predictive power (Booker 2020; Couce et al. 2024). Nevertheless, the redundancy and plasticity inherent in regulatory and metabolic networks mean that many different sequence changes can produce similar phenotypic outcomes, really complicating the mapping from genotype to fitness (Lässig et al. 2017). In sum, while significant obstacles remain, particularly in generalizing predictions across biological scales and evolutionary contexts, the integration of real-time data and mechanistic models holds promise for improving our ability to anticipate evolutionary change (Lässig et al. 2017; Booker 2020; Couce et al. 2024; Luksza and Lässig 2014).

5 Conclusions

Theoretical progress has enriched forecasting methodologies through moment-based descriptions, cumulant expansions, and traveling wave models, which collectively capture not only shifts in mean fitness but also the evolution of higher order moments and stochastic fluctuations arising from finite population sizes and random mutational events. Yet substantial challenges remain in extending predictive accuracy over longer evolutionary timescales. Complexities such as sequence-level evolution, epistatic interactions, and environmental variability introduce additional layers of uncertainty that limit deterministic forecasting. The inherently stochastic nature of evolutionary processes and the contingent sequence of mutational events impose fundamental constraints on long-term predictability.

Practical implementation of increasingly accurate forecasting will have exciting applications in applied domains like public health, where anticipating pathogen evolution can guide intervention strategies and policy decisions. The intricate interplay among mutation, selection, and genetic architecture necessitates ongoing integration of empirical data, theoretical models, and advanced statistical inference. Future progress will rely on expanding predictive frameworks to include multiple phenotypes, complex environmental interactions, and whole-genome data.

Overall, continued incremental progress is increasingly breaking through dogmatic barriers in evolutionary biology; the integration of real-time data-driven analysis, explicit modeling of evolution under varying degrees of genetic linkage, and a fitness-centered approach is reshaping evolutionary biology into a discipline capable of forecasting, and potentially influencing, evolutionary outcomes. Continued interdisciplinary collaboration will be essential to fully realize the predictive potential of evolutionary biology and translate it into effective practical applications.

6 Cross-References

▸ Data-Driven Modeling
▸ NFL Data Analytics and Predictions Using Machine Learning
▸ Parameter Estimation and Forecasting for Biased Models
▸ Quantitative Approaches in the Life Sciences
▸ Using Markov Chains in Predictive Modeling of Sports

Competing Interest Declaration The author(s) has no competing interests to declare that are relevant to the content of this manuscript.

References

Aguilar RD (2025) Leveraging the strong biasing effects of natural selection to characterize fitness effects of adaptive mutations in evolving in situ populations. Master's thesis, Univeridad Autonoma del Estado de Hidalgo, Mexico

Amir R, Evstigneev IV, Hens T, Potapova V, Schenk-Hoppé KR (2021) Evolution in pecunia. Proc Natl Acad Sci U S A 118(26):e2016514118

Autzen B (2024) The foundations of bet hedging. Biol Philos 39(6):1–14

Barton NH, Coe JB (2009) On the application of statistical physics to evolutionary biology. J Theor Biol 259(2):317–324

Becker N, Sibani P (2014) Evolution and non-equilibrium physics: a study of the Tangled Nature Model. EPL 105(1):18005

Bertram J, Gomez K, Masel J (2017) Predicting patterns of long-term adaptation and extinction with population genetics. Evolution 71(2):204–214

Bisschop K, Blankers T, Mariën J, Wortel MT, Egas M, Groot AT, Visser ME, Ellers J (2022) Population bottleneck has only marginal effect on fitness evolution and its repeatability in dioecious *Caenorhabditis elegans*. Evolution 76:1896–1904. https://doi.org/10.1111/evo.14556

Booker TR (2020) Inferring parameters of the distribution of fitness effects of new mutations when beneficial mutations are strongly advantageous and rare. G3 (Bethesda) 10(7):2317–2326

Bräutigam C, Smerlak M (2022) Diffusion approximations in population genetics and the rate of Muller's ratchet. J Theor Biol 550:111236

Cecconi F, Cencini M, Falcioni M, Vulpiani A (2012) Predicting the future from the past: An old problem from a modern perspective. Am J Phys 80(11):1001–1008

Chia N, Goldenfeld N (2011) Statistical Mechanics of Horizontal Gene Transfer in Evolutionary Ecology. J Stat Phys 142(6):1287–1301

Colato A, Fontanari JF (2001) Soluble model for the accumulation of mutations in asexual populations. Phys Rev Lett 87(23):238102

Colato A, Fontanari JF, Howard S (2005) Evolução de populações assexuadas em um sistema de presa-predador. In: Resumos, Sociedade Brasileira de Física – SBF, São Paulo

Couce A, Limdi A, Magnan M, Owen SV, Herren CM, Lenski RE, Tenaillon O, Baym M (2024) Changing fitness effects of mutations through long-term bacterial evolution. Science 383(6681):1417

Diep HT, Desgranges G (2021) Dynamics of the price behavior in stock markets: a statistical physics approach. Physica A 570(125813):125,813

Dmouj A (2006) Stock price modelling: theory and Practice. PhD thesis, Vrije Universiteit, Amsterdam

Figgins MD, Bedford T (2025) Frequency dynamics predict viral fitness, antigenic relationships and epidemic growth. medRxiv p 2024.12. 02.24318334

Fisher D, Lässig M, Shraiman B (2013) Evolutionary dynamics and statistical physics. J Stat Mech 2013(01):N01001

Fontanari JF, Colato A, Howard RS (2003) Mutation accumulation in growing asexual lineages. Phys Rev Lett 91(21):218101

Gallegos C, Chevin LM, Hodgins KA, Monro K (2025) Predicting adaptation and evolution of plasticity from temporal environmental change. Methods Ecol Evol 16(1):84–96

Gandon S, Day T, Metcalf CJE, Grenfell BT (2016) Forecasting epidemiological and evolutionary dynamics of infectious diseases. Trends Ecol Evol 31(10):776–788

Gerrish P (2001) The rhythm of microbial adaptation. Nature 413(6853):299–302

Gerrish PJ (2016) Causes of Broadly-neutralizing Antibody (BnAb) evolution, and approaches to engineering BnAb evolution through vaccination. Tech. rep., The Duke Center for HIV/AIDS Vaccine Immunology

Gerrish PJ, Hengartner N (2017) Inferring the distribution of fitness effects (DFE) of newly-arising mutations using samples taken from evolving populations in real time. In: Algorithms for computational biology. Springer International Publishing, Cham, pp 103–114

Gerrish PJ, Lenski RE (1998) The fate of competing beneficial mutations in an asexual population. Genetica 102–103(1–6):127–144
Gerrish PJ, Sniegowski PD (2011) Adding dynamical sufficiency to Fisher's fundamental theorem of natural selection. In: Simos TE, Psihoyios G, Tsitouras C, Anastassi Z (eds) Numerical analysis and applied mathematics. Springer, New York, NY
Gerrish PJ, Sniegowski PD (2012) Real time forecasting of near-future evolution. J R Soc Interface 9(74):2268–2278
Gerrish PJ, Colato A, Perelson AS, Sniegowski PD (2007) Complete genetic linkage can subvert natural selection. Proc Natl Acad Sci U S A 104(15):6266–6271
Gerrish PJ, Colato A, Sniegowski PD (2013) Genomic mutation rates that neutralize adaptive evolution and natural selection. J R Soc Interface 10(85):20130,329
Girault V, Raviart PA (2011) Finite element methods for Navier-stokes equations: theory and algorithms. Springer series in computational mathematics. Springer, Berlin
Haigh J (1978) The accumulation of deleterious genes in a population–Muller's Ratchet. Theor Popul Biol 14(2):251–267
Heilbron K, Toll-Riera M, Kojadinovic M, MacLean RC (2014) Fitness is strongly influenced by rare mutations of large effect in a microbial mutation accumulation experiment. Genetics 197(3):981–990
Koduri N, Lo AW (2021) The origin of cooperation. Proc Natl Acad Sci U S A 118(26):e2015572118
Kussell E, Vucelja M (2014) Non-equilibrium physics and evolution—adaptation, extinction, and ecology: a Key Issues review. Rep Prog Phys 77(10):102602
Levin SA, Lo AW (2021) Introduction to PNAS special issue on evolutionary models of financial markets. Proc Natl Acad Sci U S A 118(26):e2104800,118
Lind PA, Libby E, Herzog J, Rainey PB (2019) Predicting mutational routes to new adaptive phenotypes. Elife 8:e38822
Lorenz EN (1963) Deterministic nonperiodic flow. J Atmos Sci 20(2):130–141
Lässig M, Mustonen V, Walczak AM (2017) Predicting evolution. Nat Ecol Evol 1(3):77
Luksza M, Lässig M (2014) A predictive fitness model for influenza. Nature 507(7490):57–61
Lynch P (2006) The emergence of numerical weather prediction: Richardson's dream. Cambridge University Press, Cambridge
Lynch P (2016) An artist's impression of Richardson's fantastic forecast factory. Weather 71(1): 14–18
Martin G (2014) Fisher's geometrical model emerges as a property of complex integrated phenotypic networks. Genetics 197(1):237–255
Martin G, Roques L (2016) The nonstationary dynamics of fitness distributions: Asexual model with epistasis and standing variation. Genetics 204(4):1541–1558
Matsuda H, Ogita N, Sasaki A, Satō K (1992) Statistical mechanics of population: the Lattice Lotka-Volterra model. Progr Theoret Phys 88(6):1035–1049
Merleau NSC, Pénisson S, Gerrish PJ, Elena SF, Smerlak M (2021) Why are viral genomes so fragile? The bottleneck hypothesis. PLoS Comput Biol 17(7):e1009128
Miralles R, Gerrish PJ, Moya A, Elena SF (1999) Clonal interference and the evolution of RNA viruses. Science 285(5434):1745–1747
Mustonen V, Lässig M (2009) From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation. Trends Genet 25(3):111–119
Namias J (1968) Long range weather forecasting—history, current status and outlook1. Bull Am Meteorol Soc 49(5-1):438–470
Neher RA, Shraiman BI (2011) Statistical genetics and evolution of quantitative traits. Rev Mod Phys 83(4):1283–1300
Nosil P, Flaxman SM, Feder JL, Gompert Z (2020) Increasing our ability to predict contemporary evolution. Nat Commun 11(1):5592
Passagem-Santos D, Zacarias S, Perfeito L (2018) Power law fitness landscapes and their ability to predict fitness. Heredity 121(5):482–498

Perfeito L, Fernandes L, Mota C, Gordo I (2007) Adaptive mutations in bacteria: high rate and small effects. Science 317(5839):813–815

Perfeito L, Ghozzi S, Berg J, Schnetz K, Lässig M (2011) Nonlinear fitness landscape of a molecular pathway. PLoS Genet 7(7):e1002160

Peters O, Neelin JD (2006) Critical phenomena in atmospheric precipitation. Nat Phys 2(6): 393–396

Pielke RA Sr, Shen BW, Zeng X (2024) The butterfly effect: Can a butterfly in Brazil really cause a tornado in Texas? Weatherwise 77(3):14–18

Pénisson S, Sniegowski PD, Colato A, Gerrish PJ (2013) Lineage dynamics and mutation–selection balance in non-adapting asexual populations. J Stat Mech 2013(01):P01013

Pénisson S, Singh T, Sniegowski P, Gerrish P (2017) Dynamics and fate of beneficial mutations under lineage contamination by linked deleterious mutations. Genetics 205(3):1305–1318

Reeve JP (2000) Predicting long-term response to selection. Genet Res Camb 75:83

Richardson LF (1922) Weather prediction by numerical process. The University Press, Cambridge

Robson AJ, Orr HA (2021) Evolved attitudes to risk and the demand for equity. Proc Natl Acad Sci U S A 118(26):e2015569118

Rozen DE, de Visser JAGM, Gerrish PJ (2002) Fitness effects of fixed beneficial mutations in microbial populations. Curr Biol 12(12):1040–1045

Sarewitza D, Pielke Jr R (1999) Prediction in science and policy. Technol Soc 21:121–133. www.elsevier.com/locate/techsoc

Sella G, Hirsh AE (2005) The application of statistical physics to evolutionary biology. Proc Natl Acad Sci U S A 102(27):9541–9546

Shen BW, Pielke RA Sr, Zeng X, Zeng X (2023) Lorenz's view on the predictability limit of the atmosphere. Encyclopedia (Basel, 2021) 3(3):887–899

Shuman FG (1989) History of numerical weather prediction at the national meteorological center. Weather Forecast 4(3):286–296

Smerlak M (2018) Natural selection as coarsening. J Stat Phys 172(1):105–113

Smerlak M (2019) Effective potential reveals evolutionary trajectories in complex fitness landscapes. arXiv preprint arXiv:191205890. https://arxiv.org/abs/1912.05890

Smerlak M (2021) Quasi-species evolution maximizes genotypic reproductive value (not fitness or flatness). J Theor Biol 522:110,699

Smerlak M, Youssef A (2017) Limiting fitness distributions in evolutionary dynamics. J Theor Biol 416:68–80

Sniegowski PD, Gerrish PJ (2010) Beneficial mutations and the dynamics of adaptation in asexual populations. Philos Trans R Soc Lond B Biol Sci 365(1544):1255–1263

Sniegowski PD, Gerrish PJ, Lenski RE (1997) Evolution of high mutation rates in experimental populations of E. coli. Nature 387(6634):703–705

Trindade S, Perfeito L, Gordo I (2010) Rate and effects of spontaneous mutations that affect fitness in mutator Escherichia coli. Philos Trans R Soc B Biol Sci 365(1544):1177–1186

de Visser JAGM, Zeyl CW, Gerrish PJ, Blanchard JL, Lenski RE (1999) Diminishing returns from mutation supply rate in asexual populations. Science 283(5400):404–406

de Vladar HP, Barton NH (2011) The contribution of statistical physics to evolutionary biology. Trends Ecol Evol 26(8):424–432

Wortel MT, Agashe D, Bailey SF, Bank C, Bisschop K, Blankers T, Cairns J, Colizzi ES, Cusseddu D, Desai MM, van Dijk B, Egas M, Ellers J, Groot AT, Heckel DG, Johnson ML, Kraaijeveld K, Krug J, Laan L, Lässig M, Lind PA, Meijer J, Noble LM, Okasha S, Rainey PB, Rozen DE, Shitut S, Tans SJ, Tenaillon O, Teotónio H, de Visser JAGM, Visser ME, Vroomans RMA, Werner GDA, Wertheim B, Pennings PS (2023) Towards evolutionary predictions: current promises and challenges. Evol Appl 16:3–21. https://doi.org/10.1111/eva.13513

Wu T, Gao X, An F, Sun X, An H, Su Z, Gupta S, Gao J, Kurths J (2024) Predicting multiple observations in complex systems through low-dimensional embeddings. Nat Commun 15(1):2242

Zhou HJ (2016) The rock–paper–scissors game. Contemp Phys 57(2):151–163

Zhou J, McCandlish DM (2020) Minimum epistasis interpolation for sequence-function relationships. Nat Commun 11(1):1782