Vol. 20, e200034 (2023)
doi: 10.2142/biophysico.bppb-v20.0034

Biophysics and Physicobiology
https://www.jstage.jst.go.jp/browse/biophysico/

Review Articles (Invited)

Population dynamics models for various forms of adaptation

So Nakashima[1], Tetsuya J. Kobayashi[1,2,3,4]

[1] *Institute of Industrial Science, The University of Tokyo, Meguro-ku, Tokyo 153-8505, Japan*
[2] *Universal Biology Institute, The University of Tokyo, Bunkyo-ku, Tokyo 113-8654, Japan*
[3] *Department of Mathematical Informatics, Graduate School of Information Science and Technology, The University of Tokyo, Bunkyo-ku, Tokyo 113-8654, Japan*
[4] *Department of Electrical Engineering and Information Systems, Graduate School of Information Science and Technology, The University of Tokyo, Bunkyo-ku, Tokyo 113-8654, Japan*

Received May 11, 2023; Accepted August 31, 2023;
Released online in J-STAGE as advance publication September 2, 2023
Edited by Tomoaki Matsuura

Adaptability to changing environments is one of the universal characteristics of living organisms. Because individual modes of adaptation are diverse, a unified understanding of these diverse modes is essential to comprehend adaptation. Adaptations can be categorized from at least two perspectives with respect to information. One is the passivity and activity of adaptation and the other is the type of information transmission. In Darwinian natural selection, organisms are selected among randomly generated traits under which individual organisms are passive in the sense that they do not process any environmental information. On the other hand, organisms can also adapt by sensing their environment and changing their traits. This is an active adaptation in that it makes use of environmental information. In terms of information transfer, adaptation through phenotypic heterogeneity, such as bacterial bet-hedging, is intragenerational in which traits are not passed on to the next generation. In contrast, adaptation through genetic diversity is intergenerational. The theory of population dynamics enables us to unify these various modes of adaptations and their properties can be analyzed qualitatively and quantitatively using techniques from quantitative genetics and information thermodynamics. In addition, such methods can be applied to situations where organisms can learn from past experiences and pass them on from generation to generation. In this work, we introduce the unified theory of biological adaptation based on population dynamics and show its potential applications to evaluate the fitness value of information and to analyze experimental lineage tree data. Finally, we discuss future perspectives for its development. This review article is an extended version of the Japanese article in SEIBUTSU BUTSURI Vol. 57, p.287-290 (2017).

Key words: mathematical biology, natural selection, information processing, variational representation

◄ ***Significance*** ►

Population dynamics is a mathematical model to quantitatively analyze various types of adaptation. Although classical population dynamics is powerful enough to analyze passive adaptation by natural selection, it can be generalized to cover broader types of adaptation by employing techniques from other fields such as stochastic thermodynamics and information thermodynamics. For example, the variational representation allows us to calculate the population fitness analytically. Further generalization reveals a beautiful connection between the fitness value of sensing and information thermodynamics. Moreover, population dynamics is applied to analyze various types of experimental data including lineage tree data without survivorship bias.

Corresponding author: So Nakashima, Institute of Industrial Science, The University of Tokyo, Meguro-ku, Tokyo 153-8505, Japan. ORCID iD: https://orcid.org/0000-0002-5287-7475, e-mail: naaso0510@gmail.com

Introduction: Adaptation is a Fundamental Characteristic of Organisms

Organisms can adapt to their surrounding environment in various ways (Fig. 1 (a)). An example is Darwinian natural selection [1]: A population of organisms is genetically and phenotypically diverse due to mutation and stochastic gene expression; Organisms more fitted to the environment are more likely to survive and reproduce offspring than others; Consequently, such organisms dominate in the population. We can regard natural selection as a passive adaptation, since mutation generates a variety of organisms randomly rather than intendedly (Fig. 1 (b)). At the same time, many organisms, from humans to microbes, can actively adapt to the environment by sensing the state of environment and subsequently changing their behavior and phenotypes intendedly for survival and reproduction (Fig. 1 (c)) [2, 3]. Thus, there are both passive and active adaptations.

Adaptation also takes on various forms in the way how descendants inherit the traits of parents. When adaptation is driven by mutation and natural selection, traits of individual organisms are encoded genetically and transmitted to their descendants through the inheritance of genes [4]. When adaptation is achieved actively by environment sensing, the change in traits cannot be transmitted genetically to the descendants [4]. This difference in transmission affects the time scale of adaptation; Adaptation with genetic transmission is intergenerational, while adaptation without genetic transmission is mostly intragenerational (Certain phenotypic traits can be transmitted intergenerationally via epigenetic.).

These various forms of adaptation have been investigated quantitatively by using a mathematical model called population dynamics. Population dynamics was first applied to passive adaptation by natural selection and revealed the relationship between the variety of phenotypes in the population and the speed of passive adaptation. This relationship is known as Fishers fundamental theorem of natural selection and Prices equation [5,6]. Furthermore, population dynamics can predict how the changes in phenotypes or traits affect the speed of population growth [7–9]. This traditional theory of passive adaptation by natural selection has recently been integrated with information theory and stochastic thermodynamics to accommodate active adaptation into population dynamics. Furthermore, the applications of population dynamics, which were originally dominated by evolution and adaptation of higher organisms, such as animals, have now expanded to the microevolution of microbes, including persistence to antibiotics.

In this review, we overview population dynamics and its connection to both passive and active adaptations by focusing on whether they are intergenerational or intragenerational (Fig. 1 (d) and (e)). To focus on physical and biological implications, we omit a detailed derivation of the results. Refer to the following references [10–15] for the omitted derivations.

Population Dynamics of Intragenerational Adaptation (Type 1 in Fig. 1 (e))

Growth of Phenotypically Inhomogeneous Population

We first introduce population dynamics that models passive adaptation in one generation. This type of population dynamics typically aims at modeling microevolution of microbes, in which we consider the change of the frequency of the genotypes or the phenotypes in a population. The emergence or extinction of alleles, phenotypes, and species are not considered in microevolution whereas these factors are explicitly under consideration in macroevolution. See the Discussion section for the modeling of macroevolution. In a population of microbes like *E. coli*, their phenotype is inhomogeneous, even when they are genetically identical. The expression level of genes and the division times are different between cells. The difference in such phenotypic traits is not mere noise, but plays a vital role in adaptation to the environment. Due to the inhomogeneity of the traits, a part of the population may be able to survive even when the environmental conditions change drastically. This survival strategy of a population is called bet-hedging [16]. Bet-hedging is microevolution since no traits neither emerge nor extinct.

An example of bet-hedging is bacterial persistence [17–19] in which a part of microbes survives for a long time even after antibiotics exposure. We can also observe bet-hedging by using artificial gene circuits [20]. In addition, the difference in the division time of cells can be beneficial for population growth [21–23]. The inhomogeneity of the division time enables cells that happen to replicate faster than average to generate more descendants, resulting in a shortening of the average division time of the population compared to that of individuals. Thus, it accelerates the growth of the population.

Researchers have tried to understand these phenomena more quantitatively by using models of population dynamics (for bed-hedging [24–28]; for the inhomogeneity of division times [21, 22, 29].). These phenomena are typical examples of passive adaptation in a sense that the randomly generated heterogeneity of the population leads to the acceleration of population's growth. In addition,phenotypic heterogeneity is not transmitted via gene and can change drastically upon cell divisions; thus, the adaptation is mainly intragenerational (It should be noted that certain phenotypic states can be transmitted to the next generation.).

Let us mathematically model these phenomena. We focus on a population of organisms that reproduce assexually.

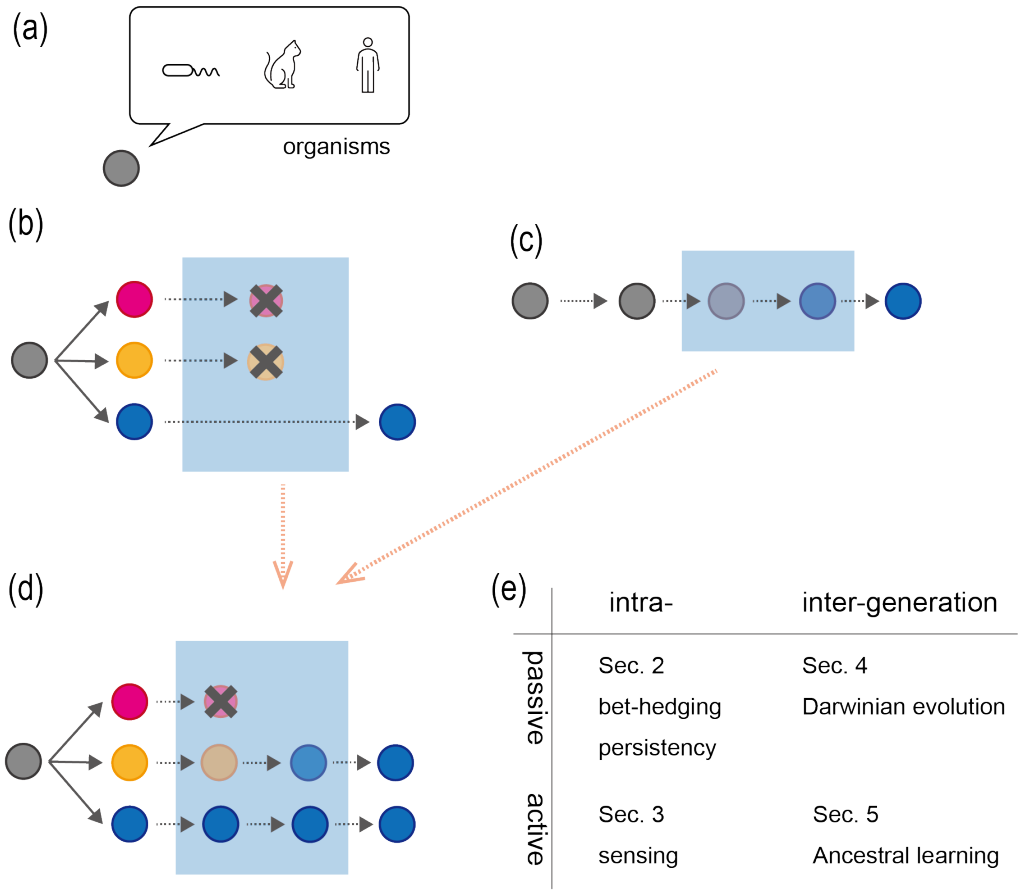

Figure 1 Schematic representation of various kinds of adaptation. (a) We consider organisms including microbes, animals, and humans in this review. (b) Schematic representation of passive adaptation. A population of organisms is phenotypically diverse, and the organisms fitted more to the environment (e.g., blue one in the panel) survive and reproduce. (c) Schematic representation of active adaptation. Organisms sense the environment and change their phenotypes (e.g., from gray to blue in the panel) so that they fit more into the environment and become more likely to survive. (d) Passive and active adaptation may occur simultaneously. An example is ancestral learning explained in Sec. 5. (e) In this review, we outline population dynamics and its connection to both passive and active adaptations one by one. We also focus on whether the adaptation is intragenerational or intergenerational.

We can apply the model to the population that reproduces sexually by focusing only on female organisms. We abstract inhomogeneous phenotypes such as gene expression level by type x. Let x be the set of possible types. We assume that the type is unchanged within one generation for simplicity. The type is assumed to change intergenerationally and is determined when an organism replicates. The probability that the type upon replication becomes x is $\pi(x)$ (Fig. 2 (a)). We can regard $\pi(x)$ as being genetically encoded, i.e., the probability to take a certain type is determined genetically. We first focus on the situation that the population is genetically homogeneous, i.e., organisms share the same $\pi(x)$. For the environmental state y, we assume that an organism with type x reproduces $e^{k(x,y)}$ daughters on average (Fig. 2 (b)). We suppose that the environmental state y_t at time t follows probability distribution $Q(y)$ independently (at the cost of theoretical complexity, we can also consider a temporally dependent environmental change [30, 31].). Under this setting, the number of organisms with type x is

$$N_{t+1}(x) = \pi(x)e^{k(x,y_{t+1})} \sum_{x' \in \mathcal{X}} N_t(x'). \tag{1}$$

This model implicitly assumes no correlation between the type of parent x' and the type of daughter x. The type

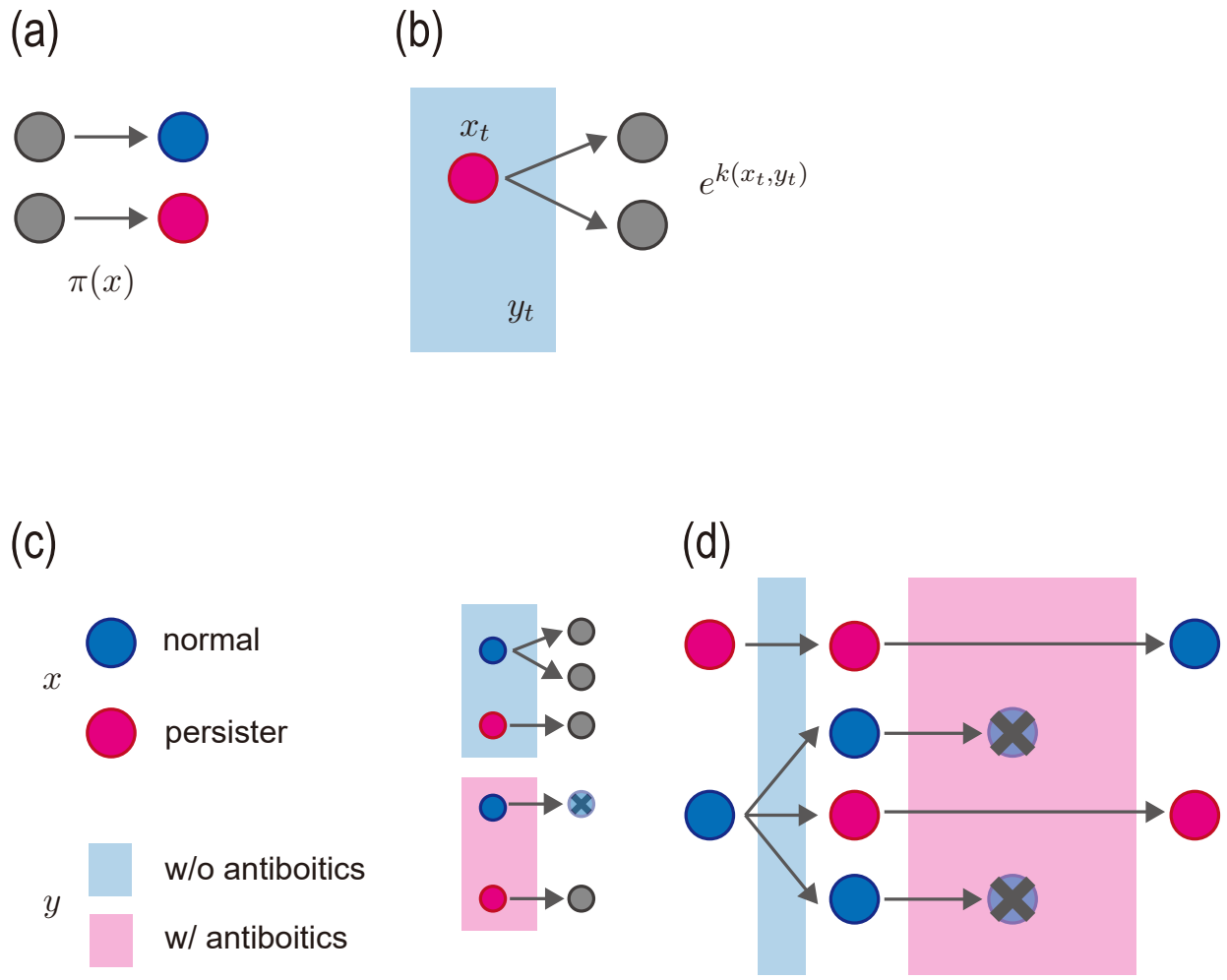

Figure 2 Schematic illustration of population dynamics model. (a) The type of an organism is stochastically determined by $\pi(x)$ at its birth. (b) An organism of type x_t produces $e^{k(x_t,y_t)}$ daughters on average when the environmental state is y_t. (c-d) An example model of persistence. (c) We consider a population of two types of cells: normal and persistent. We also consider two types of environmental states: those with and without antibiotics. Normal cells are more fitted to the environment without antibiotics than persister cells, whereas the converse holds for the environment with antibiotics. (d) By simulating the model (c), we can analyze which bet hedging strategy $\pi(x)$ is the most advantageous when environmental conditions change.

information is not transmitted and is thus intragenerational. When considering the correlation, we can generalize (1) by substituting $\pi(x)$ with a conditional probability $T(x|x')$ that satisfies $\sum_{x\in\mathcal{X}} T(x|x') = 1$ for all x'. Then, the model becomes

$$N_{t+1}(x) = \sum_{x'\in\mathcal{X}} e^{k(x,y_{t+1})} T(x|x') N_t(x'). \tag{2}$$

In this review, we mainly focus on non-correlation models (1) for simplicity. See references [10, 12] for the generalization to correlation models (2).

The speed of population growth defines population fitness. The cumulative population fitness is defined by

$$\Lambda_T(\pi; \{y_1, \ldots, y_T\}) = \log \frac{\sum_{x\in\mathcal{X}} N_t(x)}{\sum_{x\in\mathcal{X}} N_0(x)}, \tag{3}$$

and the time-averaged population fitness is defined by

$$\lambda(\pi) = \lim_{T\to\infty} \frac{1}{T}\Lambda_T(\pi; \{y_1, \ldots, y_T\}). \tag{4}$$

It is known that the time-averaged population fitness converges to the same value for almost all realization of the environmental condition y_t (Concretely, it converges to the same constant value with probability one under probability measure $\prod_{t=0}^{\infty} Q(y_t)$ [30, 31]). We call the time-averaged population fitness just population fitness in short. Population fitness intuitively means that the size of the population grows exponentially as $\sum_{x\in\mathcal{X}} N_t(x) \sim e^{\lambda(\pi)t}$. Population fitness is a natural measure of population growth owing to the following fact: If $\lambda(\pi) > \lambda(\pi')$, then the size $\sum_{x\in\mathcal{X}} N_t(x)$ of population with π is larger than that of population with π' for sure when t is large enough (when t is small, then the size of the population with smaller $\lambda(\pi)$ might be larger by chance [7, 32]).

Let us illustrate how to model a specific scenario of adaptation using a simple example of population dynamics (Fig. 2 (c)). To model the persistence phenomenon, let us consider a population of cells with two types. One type is normal x_{normal} and the other is persister $x_{\text{persister}}$. A cell becomes either normal or persister upon birth with probability $\pi(x)$ independently of the type of its parent. Normal cells can replicate faster than persisters under environmental condition y_{normal} without antibiotics. On the other hand, persister cells replicate faster under environmental condition y_{anti} with antibiotics. This situation can be modeled by choosing k so that $k(x_{\text{normal}}, y_{\text{normal}}) > k(x_{\text{persister}}, y_{\text{normal}})$ and $k(x_{\text{normal}}, y_{\text{anti}}) < k(x_{\text{persister}}, y_{\text{anti}})$. The environmental condition follows a probability distribution $Q(y)$. By simulating this model for various π and Q, we can analyze which bet-hedging strategy π is the most advantageous when the environmental condition ($k(x, y)$ and $Q(y)$) changes (Fig. 2 (d)).

Path-integral Formulation and Variational Representation of Population Fitness

To reveal the relationship between population fitness $\lambda(\pi)$ and model parameters ($\pi(x)$ and $k(x, y)$), techniques from stochastic thermodynamics have recently been employed. These techniques enable us to represent the population fitness analytically.

We first explain a path-integral expression of population fitness. In a path-integral approach [10], we consider the reproduction of organisms over several generations at once. This approach enables us to calculate the population fitness and other quantities related to the population more easily and transparently than the original model (1). Concretely, a path-integral expression of population fitness is following:

$$\Lambda(\pi; Y) = \log\langle e^{k[X,Y]}\rangle_{P_F[X]}. \tag{5}$$

Here, $X = \{x_1, \ldots, x_T\}$ is the path (history) of phenotypes of organisms over several generations and $Y = \{y_1, \ldots, y_T\}$ is the path of environmental states history. The quantity $P_F[X] = \prod_{t=1}^{T} \pi(x_t)$ is the probability that the path of types is X when we ignore the effect of population growth. We call $P_F[X]$ a forward probability. The quantity $k[X, Y] := \sum_{t=1}^{T} k(x_t, y_t)$ represents how many offspring are generated along a lineage of organisms whose path of types is X. The quantity $\langle F[X]\rangle_{p(X)} := \mathrm{E}_{p(X)}[F[X]]$ is an abbreviation of expectation of $F[X]$ with respect to $P[X]$. The relationship (5) is useful because it holds for a broad class of models. For example, the path-integral expression holds for (2).

We can formally regard (5) as a free energy by identifying $k(x, y)$ with negative energy. Therefore, we can rewrite (5) in a variational form.

$$\Lambda(\pi; Y) = \max_{P[X]} \left[\sum_X k[X, Y]P[X] - \mathrm{KL}[P[X]|P_F[X]] \right], \tag{6}$$

where $\mathrm{KL}[Z[X], W[X]]$ is the Kullbuck-Leibler divergence of two probability distributions $Z[X]$ and $W[X]$ defined by $\mathrm{KL}[Z[X], W[X]] = \sum_X Z[X] \log \frac{Z[X]}{W[X]}$. The Kullbuck-Leibler divergence measures distance between two probability distributions. When $W[X]$ is a uniform distribution, the Kullbuck-Leibler divergence is reduced to with the negative entropy. In (6), maximization is taken over all path probability distribution $P[X]$. The explicit form of the maximizer $P_B[X]$ of (6) is obtained as

$$P_B[X|Y] :\propto e^{k[X,Y]} P_F[X]. \tag{7}$$

The probability distribution $P_B[X]$ is called a backward probability [10, 33]. In (1), we can calculate the backward probability more explicitly as

$$P_B[X|Y] = \prod_{t=1}^{T} \pi_B(x_t, y_t), \tag{8}$$

$$\pi_B(x, y) :\propto e^{k(x,y)} \pi(x). \tag{9}$$

By using the backward probability, we can rewrite (6) as

$$\Lambda(\pi; Y) = \sum_X k[X, Y] P_B[X] - \mathrm{KL}[P_B[X]|P_F[X]]. \tag{10}$$

The path-level probability $P_F[X]$ and $P_B[X]$ are not mere tools of calculation. We can interpret and connect these theoretical quantities to experimentally measurable ones. The forward probability $P_F[X]$ is the probability that we observe a path X of types when we sample a lineage of organisms by the following rule: we observe a lineage chronologically by choosing one of the daughters uniformly at random at each division. We can experimentally sample $P_F[X]$ using microfluidic devices such as a mother machine [34]. The backward probability $P_B[X]$ is the probability of obtaining path X when we sample a lineage according to the following rule: We first choose an organism at the end of the experiment uniformly at random; then we follow the ancestral lineage of the chosen organism backward. Since a lineage generating more offspring has more descendants at the end of the experiment, such a lineage is more likely to appear in $P_B[X]$ than in $P_F[X]$. In other words, population growth and natural selection biases $P_B[X]$ from $P_F[X]$. Thus, we can measure the impact of selection biases by $\mathrm{KL}[P_B[X]|P_F[X]]$. We can experimentally determine $P_B[X]$ by sampling cells using the dynamics cytometer [23]. In the above interpretation, the right-hand side of (10) is the difference between the expected population growth of the lineage sampled by backward probability (first term) and the magnitude of biases induced by natural selection (second term). The latter term is useful for analyzing and estimating selection pressure from experimental data [35].

Linear Response of Population Fitness

An application of variational representation (6) is to quantitatively predict how much population fitness changes when the average number $e^{k(x,y)}$ of offspring or the type distribution $\pi(x)$ changes. We can calculate the derivative of the population fitness via (10) since $P_B[X]$ is the maximizer of (6). The derivative is then

$$\delta\Lambda[\pi; Y] = \sum_{X,Y} P_B[X|Y] \delta k[X, Y] + \sum_X \frac{P_B[X]}{P_F[X]} \delta P_F[X]. \tag{11}$$

This equation implies that the backward probability determines the sensitivity (i.e., the derivative) of the population fitness to the change in model parameters. This relationship has been proven for a broad class of models [10, 12]. By this equation and the constraint $\sum_x \pi(x) = 1$, the population fitness attains the maximum value when the following consistency condition holds:

$$\pi(x) = \langle \pi_B(x|y) \rangle_{Q(y)}. \tag{12}$$

This equation intuitively means that natural selection does not marginally affect the type distribution if it is optimally adapted to the environment. The consistency condition was proven in another context without a backward probability [36].

Continuous-Time Population Dynamics

In previous sections, we used discrete-time population dynamics, in which the time axis is the generation time. Discrete-time models are easy to analyze theoretically. However, discrete-time models implicitly assume that the replication of organisms is synchronized, which is not always the case. For non-synchronized replication, we must use continuous-time population dynamics, in which the time axis is real time. The continuous-time model enables us to discuss the effect of inhomogeneity of the division time to population growth [21, 23].

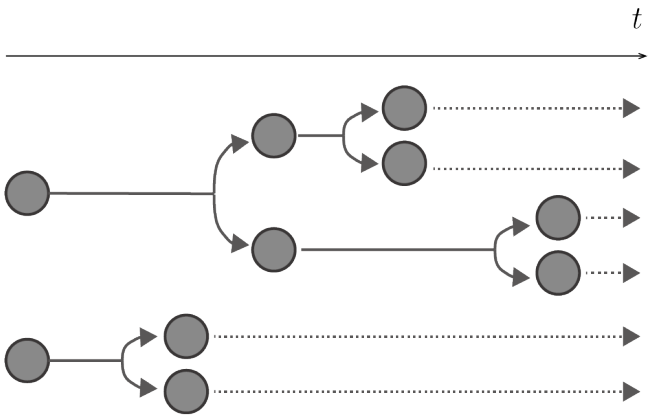

Figure 3 Schematic representation of lineage trees. Cell division is represented by the forks of arrows. The division times of cells are different, and the divisions are not synchronized.

An example of continuous-time models is age-structured population dynamics (ASPD). ASPD considers the time duration from the birth of an organism to its replication. When time length a passes since the birth of an organism, we say that the age of the organism is a. We suppose that the probability that an organism with age a and type x reproduces in an infinitesimal time duration $\mathrm{d}t$ is $r(a,x)\mathrm{d}t$. We assume that an organism always generates two offspring, and their types do not correlate with that of the parent for simplicity. In this setting, the number of organisms with age a and type x at time t follows McKendric-von Foerster equation [37, 38] (See [12] for the derivation.).

$$\frac{\partial}{\partial t}N_t(a,x) = -\frac{\partial}{\partial a}N_t(a,x) - r(a,x)N_t(a,x), \tag{13}$$

$$N_t(0,x) = 2\pi(x)\int_0^\infty \mathrm{d}a \sum_{x'\in\mathcal{X}} r(a,x')N_t(a,x'). \tag{14}$$

Here, we omit mortality term for simplicity. (When we consider the correlation of the types of a parent and daughters, the integrand in (14) is more complicated.). We can define population fitness also for ASPD, which admits a generalization of path-integral representation (5) and variational representation (6). This fact examplifies that path-integral representation and variational representation hold for a broad class of models without depending on their details.

Let us introduce two applications of ASPD. The first application is correction of a bias in experimental data of growing population. In experiments of growing population, we sometimes obtain lineage tree data like Fig. 3. For example, the dynamics cytometer produce such data. A natural statistics computed from such data is the average of some quantities over a lineage tree. An example is an average of division time over a lineage tree and can be used to measure the speed of population growth. However, it has been known that the average over lineage tree differs from the true value due to survivorship bias. For example, a naive estimator of $r(a,x)$ is

$$\hat{r}(a,x) := (\text{count of cells with age } a \text{ and type } x \text{ that divides in infinitesimal time } \mathrm{d}t)/\mathrm{d}t. \tag{15}$$

However, it is known that $\hat{r}(a,x) > r(a,x)$ in general.

The cause of the survivorship bias is that lineages with more offspring are overrepresented more in the lineage tree data. For example, an organism with shorter division time by chance has more descendants at the end of the experiment than other organisms, resulting in the overrepresentation of organisms with shorter division time in lineage tree data. The authors and other groups [14, 39–41] have proven by analyzing this bias via ASPD that

$$\hat{r}(a,x) = r(a,x) + \lambda(\pi). \tag{16}$$

This relationship enables us to correct for the survivorship bias by using a refined estimator $r_{\mathrm{refined}}(a,x) := \hat{r}(a,x) - \lambda(\pi)$. The application of this correction method to lineage tree data is our Lineage EM algorithm (LEM) [14]. LEM targets lineage tree data and estimates the unknown type x and model parameters π or r without survivorship bias.

The second application is the effect of inhomogeneous division time to population growth. Equation (16) means that

the division rate $\hat{r}(a,x)$ at the population level is higher than that $r(a,x)$ at an individual level. In other words, ASPD reveals why inhomogeneity of division time accelerates reproduction at the population level.

Active Adaptation by Sensing Environment (Type 2 in Fig. 1 (e))

In the previous sections, we have focused on the situation where type x is passively determined by genetically encoded information: Organisms do not change or modulate their phenotype x intendedly to adapt to the environment. However, most organisms actively change their traits and behaviors to adapt to the environment. Animals can make decisions. Bacteria can sense and move to or escape from chemical substances like attractants and repellants. What is the fitness value of these active adaptations? How much do these active adaptations accelerate population growth?

Researchers [11,42–45] have studied these questions by generalizing the population dynamics model (1). In the generalized model, organisms can sense the environment and obtains environmental signal z. For example, bacteria can sense the concentration y of a chemical substance in environment by a noisy sensor as $z = y+n$, where n is the noise. More generally, we suppose that information z is correlated with environmental state y as conditional probability $p(z|y)$. Organisms can change their type x using the environmental signal z. Thus, the type distribution depends on z and is written as $\pi(x|z)$. We can define population fitness in a way similar to (4) in this model.

The fitness value of sensing can be measured as the increase of population fitness by sensing the environment and subsequent switching of type. We first define the maximum population fitness without sensing as

$$\lambda^* := \max_{\pi(x)} \lambda(\pi). \tag{17}$$

We similarly define the maximum population fitness with sensing as

$$\lambda^\dagger := \max_{\pi(x|z)} \lambda(\pi). \tag{18}$$

By using consistency condition (12), we can prove the following relationship:

$$\lambda^\dagger - \lambda^* \leq I[y,z], \tag{19}$$

where $I[y,z]$ is the mutual information between environmental states y and environmental signal z defined as

$$I[y,z] = \mathrm{KL}[Q(y)p(z|y)|Q(y)p(z)]. \tag{20}$$

See [11] for the derivation. The relationship means that the fitness value of sensing is bounded by the mutual information between sensed information z and environment y. The equation connects the fitness value of active adaptation to information theory. We can interpret the equation as the generalization of the relationship between information and energy in information thermodynamics, which is often symbolized as Maxwells demon [46].

Passive Adaptation in the Timescale of Evolution (Type 3 in Fig. 1 (e))

We have outlined the intragenerational adaptation of a phenotypically homogeneous but genetically homogeneous population. When we consider adaptation over several generations, the genotypes in the population become inhomogeneous because of mutation. This additional genetic variety causes different kinds of Darwinian evolution and adaptation than those in previous sections.

We can model this situation as the population of organisms having a variety of π. Recall that we assumed so far that π is common in the population, since the population is genetically homogeneous (Type 1). Our previous modeling considered that the homogeneous genotype determines π and phenotype x of each organism is generated stochastically by following π. Therefore, we can model the genetic change by mutation as the change in π and the genetic inhomogeneity corresponds to the variety of π in the population.

Let us explain the intergenerational adaptation under this model. We suppose that the environmental state y is constant for simplicity and abbreviate $k(x,y)$ as $k(x)$. The type distribution of a parent and the daughter organisms may differ due to mutation. Let $L(\pi|\pi')$ be the conditional probability that a daughter's type distribution is π when the parent distribution is π'. The conditional probability $L(\pi|\pi')$ satisfies $\sum_{\pi'} L(\pi|\pi') = 1$. Under this setting, the number of the organisms with

π and x is

$$N_t(\pi, x) = \sum_{\pi', x'} e^{k(x)} \pi(x) L(\pi|\pi') N_{t-1}(\pi', x'). \tag{21}$$

By taking summation over x, we can obtain the number $N_t(\pi)$ of organisms with π in the population as

$$N_t(\pi) = \sum_{x \in \mathcal{X}} N_t(\pi, x) = \sum_{\pi'} \left[\sum_{x \in \mathcal{X}} e^{k(x)} \pi(x) \right] L(\pi|\pi') N_{t-1}(\pi'). \tag{22}$$

The equation is equivalent to (2) by identifying $\sum_{x \in \mathcal{X}} e^{k(x)} \pi(x)$ in the above equation with $e^{k(x)}$ in (2) . This fact means that the model (1) is applicable to analyze both intragenerational and intergenerational adaptation. For later discussion, we define

$$e^{k(\pi)} := \sum_{x \in \mathcal{X}} e^{k(x)} \pi(x). \tag{23}$$

We can predict the speed of intergenerational adaptation by Fishers fundamental theorem (FF-thm) of natural selection [5]. The theorem assumes that the time scale of mutation is very slow and negligible in several generations, for simplicity. Mathematically, the assumption is that $L(\pi|\pi') = \delta_{\pi'}(\pi)$, where $\delta_{\pi'}$ is the delta function. By substituting the delta function in (22), we have

$$N_t(\pi) = e^{k(\pi)} N_{t-1}(\pi). \tag{24}$$

Fishers fundamental theorem measures the speed of adaptation not by population fitness but by mean fitness. To define mean fitness, we first introduce the fraction of organisms with π at time t as

$$p_t(\pi) = \frac{N_t(\pi)}{N_t}, \tag{25}$$

where

$$N_t = \sum_{\pi} N_t(\pi). \tag{26}$$

The mean fitness is defined by

$$m_t = \langle e^{k(\pi)} \rangle_{p_t(\pi)}. \tag{27}$$

The mean fitness increases as the population evolves by following (24). By direct calculation, the increase is

$$m_t - m_{t-1} = V_{p_{t-1}(\pi)}[e^{k(\pi)}], \tag{28}$$

where $V_p[X]$ is the variance of a random variable X whose distribution is p. The equation is called Fisher's fundamental theorem of natural selection. It means that the mean fitness increases faster when the population has a wider variety of genotypes. The FF-thm was generalized to the situations with mutation and recombination [6,47,48].

We can derive a similar equation for population fitness. In model (24), the population fitness is defined by

$$\lambda_t := \log \frac{N_t}{N_{t-1}}. \tag{29}$$

The FF-thm is generalized as

$$\lambda_t - \lambda_{t-1} = \log \mathrm{Var}_{p_{t-1}}[e^{k(\pi)}]. \tag{30}$$

Here, $\log\mathrm{Var}_{p_{t-1}}[e^{k(\pi)}]$ is a generalization of variance called log-variance. The log-variance is defined via log-covariance as

$$\log\mathrm{Cov}_p[X,Y] := \log\frac{\langle XY\rangle_p}{\langle X\rangle_p\langle Y\rangle_p}, \tag{31}$$

$$\log\mathrm{Var}_p[X] := \log\mathrm{Cov}[X,X]. \tag{32}$$

The generalized FF-thm says that the speed of evolution is faster when the variety of population is greater. Thus, the generalized FF-thm has an interpretation similar to the original one in this sense.

Active Adaptation and Evolution: Acceleration of Evolutionary Process by Ancestral Learning (Type 4 in Fig. 1 (e))

Traditional evolutionary biology assumes that mutation does not depend on the phenotype of each organism. This assumption comes from the fact that Darwins theory of evolution assumes that descendants cannot genetically inherit acquired traits from their parents. This assumption is incorporated into the model (22) in the form that the change $L(\pi|\pi')$ of the type distribution does not depend on the type of parent x'. However, there is accumulating evidence that organisms can transmit what they experienced to descendants in non-genetic ways. Humans can transmit their acquired knowledge to the next generation as culture. Microbes or other organisms can transmit their phenotype as epigenetic states.

Is there any interesting phenomenon caused by the dependence of $L(\pi|\pi')$ on its phenotype x? We are particularly interested in the phenomena that such a dependence has fitness value: population fitness increases faster with the dependence than without it. For example, let us consider the situation where an organism tends to have the same type as its parent. If the parent of type x survives and generates its daughters, the type x is likely to contribute to survival and reproduction. Therefore, succession of the type x to descendants might be beneficial to the descendants for adapting current environment. We have some phenomena that can be interpreted as this kind of succession of type. We can relate cultural transmission of human beings to such dependence since the transmission of what parents learned might increase children's fitness if they follow what the parents did. In addition, microbes might learn intergenerationally by storing what they learned in epigenetic states, and descendants change their phenotype depending on the epigenetic information. However, it is not clear whether this simple speculation is indeed true or not. We need more detailed and quantitative investigations via mathematical modeling or simulation. Based on the intuitive speculation above, we regard the change $L(\pi|\pi', x')$ of type distribution depending on the type of parent x' as learning in a broader sense and investigate its fitness value below.

Since the relationship between learning and evolution has attracted attention only recently, we have two unsolved questions. First, we do not know what kind of learning can accelerate adaptation. To accelerate evolution, organisms must learn how to maximize population fitness, which is a property of the population, not individuals. It is unknown whether individual organisms that have no direct access to the properties of the entire population can actually maximize population fitness. Second, we do not know how to predict and interpret the acceleration of adaptation quantitatively.

For the first question, Xue and Leibler [49] numerically showed that, by a learning rule that mimics the type of parent, organisms can attain an optimal strategy. In addition, learning was also considered in the context of behavioral biology [50,51]. A similar learning rule to that of Xue and Leibler was used to calculate optimal population fitness [50,51]. In behavioral biology, researchers try to explain the behaviors of organisms as the maximization of individual fitness. When we consider fitness not from an individual but from a populational point of view, we must consider the variety of behaviors in the population and explain the distribution of behaviors in terms of maximization of population fitness. In [15], we introduced a generalized learning rule of Xue and Leiblers one called ancestral learning (Fig. 4), which we consider in this article. Ancestral learning updates π as follows:

$$\pi_{\mathrm{new}} \leftarrow \frac{1}{\tau}\sum_{t=1}^{\tau}\delta_{x_t,x}. \tag{33}$$

Here, x_t is the type of ancestors before t generations and τ is a hyperparameter that determines the interval of learning. In this learning rule, an organism first uses a strategy π for τ generations. The organism then updates the strategy to the frequency of ancestors' type up to τ generations. In this sense, organisms mimic the types of ancestors in ancestral learning.

We demonstrated via numeral simulation that ancestral learning accelerates evolution (Fig. 5). We simulated two populations that consist of organisms with ancestral learning and with mere mutation. We can observe that population

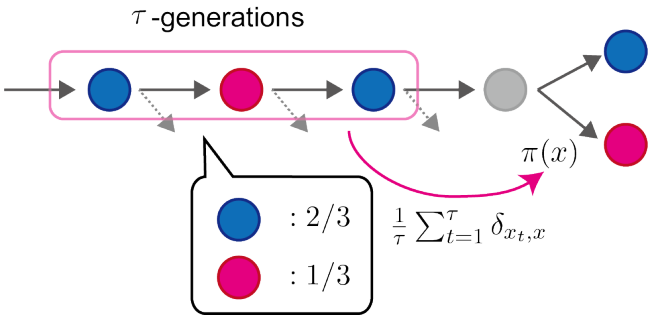

Figure 4 Schematic representation of ancestral learning. After using a bet-hedging strategy $\pi(x)$ for τ generation, an organism updates its strategy to the frequency of ancestors types up to τ generations.

fitness increases faster in the population with ancestral learning than with mutation. See [15] for details of the setting of the numerical experiment.

We then theoretically showed that the type information of the ancestors is sufficient for individuals to estimate the gradient of population fitness, i.e., individual organisms can learn how to behave to increase population fitness only from their individual ancestors behaviors. In particular, we rigorously proved that ancestral learning accelerates evolution via the relationship between ancestral learning and stochastic gradient descent. This proof used the variational representation (6) of population fitness.

For the second question, we generalized FF-thm and quantified the acceleration of evolution by ancestral learning [15]. We define the fitness difference before and after ancestral learning by

$$\Delta := \lambda(\pi_{\text{new}}) - \lambda(\pi_{\text{old}}), \tag{34}$$

where π_{new} and π_{old} are type distributions before and after learning. When the environment is constant, we can prove the following relationship by a similar argument as (30):

$$\Delta = \log \text{Var}[e^{k(\pi)}]. \tag{35}$$

Intuitively, it means that the acceleration of evolution by ancestral learning is large when the strategy is explorative and tries many types. When the environment is not constant, we can prove the following generalized relationship.

$$\Delta = \left\langle \log \text{Cov}_{\pi_{\text{old}}}\left[e^{k(x,y)}, e^{k(x,y')}\right]\right\rangle_{Q(y)Q(y')} + \text{KL}[Q(y)Q(y')|Q(y'|y)Q(y)], \tag{36}$$

$$Q(y'|y) := \sum_{x\in\mathcal{X}} e^{k(x,y)}\pi_B(x|y')Q(y'). \tag{37}$$

We can connect each of two terms in (36) to specialist strategy and generalist strategy. The first term is larger when the environmental states are more similar. In such a situation, the optimal strategy concentrates on the type x that fits best to all of the similar environmental states. Thus, the first term corresponds to the gain in population fitness by acquiring the specialist strategy. Conversely, the second term is larger when the environmental states are more dissimilar. In such a situation, the optimal strategy is usually bet-hedging. Therefore, the second term corresponds to the gain of population fitness by acquiring the generalist strategy. In this sense, the generalized FF-thm is useful for understanding under which environmental conditions ancestral learning can be beneficial.

Discussion

In this article, we review population dynamics models and their application to adaptation. Population dynamics have been generalized to many timescales (intergenerational and intragenerational) as well as to passive and active adaptation. The generalizations were achieved by incorporating techniques from other fields, including stochastic thermodynamics and information theory. Now, population dynamics has been related not only to evolutionary biology, but also to biophysics, neuroscience, and other fields. This connection accelerates the development of each field by importing knowledge from

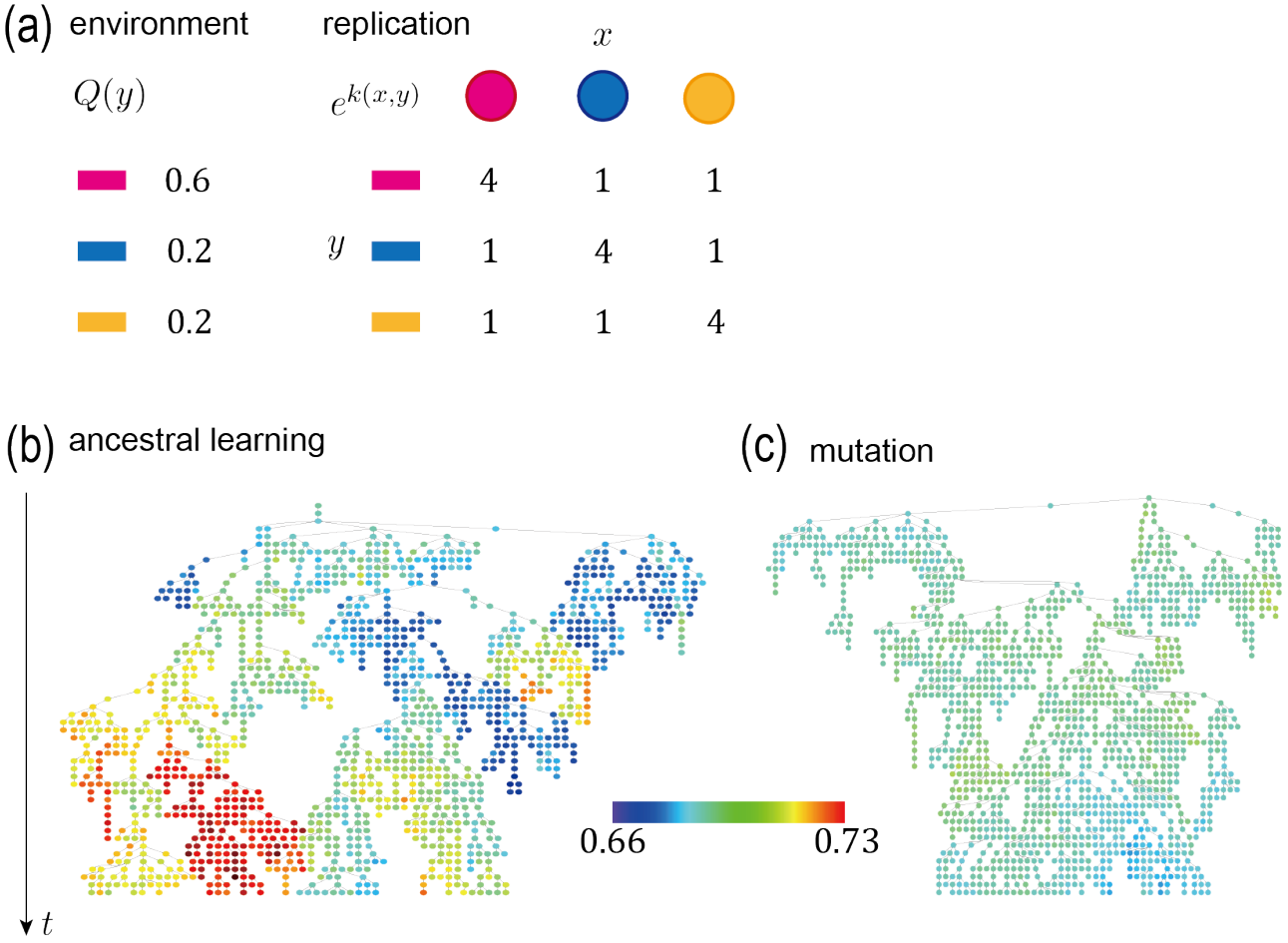

Figure 5 Numerical experiments of ancestral learning. (a) The parameters of the model. There are three environmental states. We consider a population of organisms of three types. An organism reproduces the largest number of daughters when the environmental state and the type match. (b-c) Numerical simulation of lineage trees of populations of cells with ancestral learning and with mere mutation. Each dot represents an organism. A parent and its daughter are connected by a line. The colors of the dots represent the population fitness of each organism. We can observe that the population fitness increases faster in the population with ancestral learning than only with mutation.

others. In addition, recent development in quantitative biology enables us to analyze experimental data via population dynamics and to verify the concepts of population dynamics reviewed in this article. We believe that population dynamics continues to develop by interacting with other theoretical and experimental fields. One interesting direction of further generalization is the incorporation of macroevolution. The models in this article can treat neither extinction nor emergence of phenotypes. Extinction has been investigated using stochastic population dynamics, for example, the Write-Fisher model [52]. Considering thermodynamics techniques in such models might be interesting. However, the emergence of new phenotypes is difficult to formalize as mathematical models, although some prior works have attempted to model it [53–55]. Even though the formalization is difficult, it is of course interesting and valuable not only in evolutionary biology but also in optimization as a design principle of mutation function in evolutionary algorithms.

Conflict of Interest

We have nothing to declare for this review.

Author Contributions

SN and TJK prepared the manuscripts.

Data Availability

This manuscript does not report any data. The source code of numerical simulation is available in [15].

Acknowledgement

This research is supported by JSPS KAKENHI Grants No. 19H05799, and by JST CREST No. JPMJCR2011 and No. JPMJCR1927.

References

[1] Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V. Campbell biology (Pearson, New York, 2016).
[2] McFarland, D. J. Decision making in animals. Nature 269, 15–21 (1977). https://doi.org/10.1038/269015a0
[3] Murray, J. D. Mathematical biology: I. an introduction (Springer, New York, 2002).
[4] Ridley, M. Evolution (Wiley-Blackwell, Malden, 2003).
[5] Fisher, R. A. The genetical theory of natural selection (The Clarendo Pressn, Oxford, 1930).
[6] Price, G. R. Fisher's fundamental theorem made clear. Ann. Hum. Genet. 36, 129–140 (1972). https://doi.org/10.1111/j.1469-1809.1972.tb00764.x
[7] Lewontin, R. C. Cohen, D. On population growth in a randomly varying environment. Proc. Natl. Acad. Sci. U.S.A. 62, 1056–1060 (1969). https://doi.org/10.1073/pnas.62.4.1056
[8] Arnold, L., Gundlach, V. M., Demetrius, L. Evolutionary formalism for products of positive random matrices. Ann. Appl. Probab. 4, 859 – 901 (1994). https://doi.org/10.1214/aoap/1177004975
[9] Hermisson, J., Redner, O., Wagner, H., Baake, E. Mutationselection balance: Ancestry, load, and maximum principle. Theor. Popul. Biol. 62, 9–46 (2002). https://doi.org/10.1006/tpbi.2002.1582
[10] Sughiyama, Y., Kobayashi, T. J., Tsumura, K., Aihara, K. Pathwise thermodynamic structure in population dynamics. Phys. Rev. E 91, 032120 (2015). https://doi.org/10.1103/PhysRevE.91.032120
[11] Kobayashi, T. J. Sughiyama, Y. Fluctuation relations of fitness and information in population dynamics. Phys. Rev. Lett. 115, 238102 (2015). https://doi.org/10.1103/PhysRevLett.115.238102
[12] Sughiyama, Y., Nakashima, S., Kobayashi, T. J. Fitness response relation of a multitype age-structured population dynamics. Phys. Rev. E 99, 012413 (2019). https://doi.org/10.1103/PhysRevE.99.012413
[13] Kobayashi, T. J. Sughiyama, Y. Fitness gain of individually sensed information by cells. Entropy 21, 1002 (2019). https://doi.org/10.3390/e21101002
[14] Nakashima, S., Sughiyama, Y., Kobayashi, T. J. Lineage EM algorithm for inferring latent states from cellular lineage trees. Bioinformatics 36, 2829–2838 (2020). https://doi.org/10.1093/bioinformatics/btaa040
[15] Nakashima, S. Kobayashi, T. J. Acceleration of evolutionary processes by learning and extended Fisher's fundamental theorem. Phys. Rev. Res. 4, 013069 (2022). https://doi.org/10.1103/PhysRevResearch.4.013069
[16] Seger, J. What is bet-hedging? Oxford Surveys in Evolutionary Biology 4, 182–211 (1987).
[17] Bigger, J. W. Treatment of staphyloeoeeal infections with penicillin by intermittent sterilisation. Lancet 497–500 (1944). https://doi.org/10.1016/s0140-6736(00)74210-3
[18] Balaban, N. Q., Merrin, J., Chait, R., Kowalik, L., Leibler, S. Bacterial persistence as a phenotypic switch. Science 305, 1622–1625 (2004). https://doi.org/10.1126/science.1099390
[19] Wakamoto, Y., Dhar, N., Chait, R., Schneider, K., Signorino-Gelo, F., Leibler, S., et al. Dynamic persistence of antibiotic-stressed mycobacteria. Science 339, 91–95 (2013). https://doi.org/10.1126/science.1229858
[20] Acar, M., Mettetal, J. T., van Oudenaarden, A. Stochastic switching as a survival strategy in fluctuating environments. Nat. Genet. 40, 471–475 (2008). https://doi.org/10.1038/ng.110
[21] Powell, E. O. Growth rate and generation time of bacteria, with special reference to continuous culture. Microbiology 15, 492–511 (1956). https://doi.org/10.1099/00221287-15-3-492
[22] Wakamoto, Y., Grosberg, A. Y., Kussell, E. Optimal lineage principle for age-structured populations. Evolution 66, 115–134 (2012). https://doi.org/10.1111/j.1558-5646.2011.01418.x
[23] Hashimoto, M., Nozoe, T., Nakaoka, H., Okura, R., Akiyoshi, S., Kaneko, K., et al. Noise-driven growth rate gain in clonal cellular populations. Proc. Natl. Acad. Sci. U.S.A. 113, 3251–3256 (2016).

https://doi.org/10.1073/pnas.1519412113
[24] Real, L. A. Fitness, uncertainty, and the role of diversification in evolution and behavior. Am. Nat. 115, 623–638 (1980). https://doi.org/10.1086/283588
[25] Haccou, P. Iwasa, Y. Optimal mixed strategies in stochastic environments. Theor. Popul. Biol. 47, 212–243 (1995). https://doi.org/10.1006/tpbi.1995.1009
[26] Wilbur, H. M. Rudolf, V. H. W. Lifehistory evolution in uncertain environments: Bet hedging in time. Am. Nat. 168, 398–411 (2006). https://doi.org/10.1086/506258
[27] Skanata, A. Kussell, E. Evolutionary phase transitions in random environments. Phys. Rev. Lett. 117, 038104 (2016). https://doi.org/10.1103/PhysRevLett.117.038104
[28] Hufton, P. G., Lin, Y. T., Galla, T. Phenotypic switching of populations of cells in a stochastic environment. J. Stat. Mech. 2018, 023501 (2018). https://doi.org/10.1088/1742-5468/aaa78e
[29] Lebowitz, J. L. Rubinow, S. I. A theory for the age and generation time distribution of a microbial population. J. Math. Biol. 1, 17–36 (1974). https://doi.org/10.1007/BF02339486
[30] Seppalainen, T. Large deviations for markov chains with random transitions. Ann. Probab. 22, 713–748 (1994). https://doi.org/10.1214/aop/1176988727
[31] Kifer, Y. Perron-Frobenius theorem, large deviations, and random perturbations in random environments. Math. Z. 222, 677–698 (1996). https://doi.org/10.1007/BF02621888
[32] Kelly, J. L. A new interpretation of information rate. The Bell System Technical Journal 35, 917–926 (1956). https://doi.org/10.1002/j.1538-7305.1956.tb03809.x
[33] Baake, E. Georgii, H.-O. Mutation, selection, and ancestry in branching models: A variational approach. J. Math. Biol. 54, 257–303 (2007). https://doi.org/10.1007/s00285-006-0039-5
[34] Wang, P., Robert, L., Pelletier, J., Dang, W. L., Taddei, F., Wright, A., et al. Robust growth of escherichia coli. Curr. Biol. 20, 1099 – 1103 (2010). https://doi.org/10.1016/j.cub.2010.04.045
[35] Yamauchi, S., Nozoe, T., Okura, R., Kussell, E., Wakamoto, Y. A unified framework for measuring selection on cellular lineages and traits. eLife 11, e72299 (2022). https://doi.org/10.7554/eLife.72299
[36] McNamara, J. M., Webb, J. N., Collins, E. J. Dynamic optimization in fluctuating environments. Proc. Biol. Sci. 261, 279–284 (1995). https://doi.org/10.1098/rspb.1995.0148
[37] M'Kendrick, A. G. Applications of mathematics to medical problems. Proceedings of the Royal Society of Edinburgh 44, 98130 (1925). https://doi.org/10.1017/S0013091500034428
[38] von Foerster, H. Some remarks on changing populations. in the kinetics of cellular proliferation. (Stohlman, J. F. ed.) pp. 382–407 (Grune and Stratton, New York, 1959).
[39] Marguet, A. Uniform sampling in a structured branching population. Bernoulli 25, 2649 – 2695 (2019). https://doi.org/10.3150/18-BEJ1066
[40] Marguet, A. A law of large numbers for branching Markov processes by the ergodicity of ancestral lineages. ESAIM Probab. Stat. 23, 638–661 (2019). https://doi.org/10.1051/ps/2018029
[41] García-García, R., Genthon, A., Lacoste, D. Linking lineage and population observables in biological branching processes. Phys. Rev. E 99, 042413 (2019). https://doi.org/10.1103/PhysRevE.99.042413
[42] Donaldson-Matasci, M. C., Bergstrom, C. T., Lachmann, M. The fitness value of information. Oikos 119, 219–230 (2010). https://doi.org/10.1111/j.1600-0706.2009.17781.x
[43] Rivoire, O. Leibler, S. The value of information for populations in varying environments. J. Stat. Phys. 142, 1124–1166 (2011). https://doi.org/10.1007/s10955-011-0166-2
[44] Kussell, E. Leibler, S. Phenotypic diversity, population growth, and information in fluctuating environments. Science 309, 2075–2078 (2005). https://doi.org/10.1126/science.1114383
[45] Xue, B. Leibler, S. Benefits of phenotypic plasticity for population growth in varying environments. Proc. Natl. Acad. Sci. U.S.A. 115, 12745–12750 (2018). https://doi.org/10.1073/pnas.1813447115
[46] Sagawa, T. Thermodynamics of information processing in small systems. Progress of Theoretical Physics 127, 1–56 (2012). https://doi.org/10.1143/PTP.127.1
[47] Frank, S. A. Natural selection. V. How to read the fundamental equations of evolutionary change in terms of information theory. J. Evol. Biol. 25, 2377–2396 (2012). https://doi.org/10.1111/jeb.12010
[48] Iwasa, Y. Free fitness that always increases in evolution. J. Theor. Biol. 135, 265–281 (1988). https://doi.org/10.1016/s0022-5193(88)80243-1
[49] Xue, B. Leibler, S. Evolutionary learning of adaptation to varying environments through a transgenerational feedback. Proc. Natl. Acad. Sci. U.S.A. 113, 11266–11271 (2016). https://doi.org/10.1073/pnas.1608756113
[50] Collins, E. J. McNamara, J. M. Finite-horizon dynamic optimisation when the terminal reward is a concave functional

of the distribution of the final state. Adv. Appl. Probab. 30, 122136 (1998). https://doi.org/10.1239/aap/1035227995
[51] McNamara, J. M., Webb, J. N., Collins, E., Székely, T., Houston, A. I. A general technique for computing evolutionarily stable strategies based on errors in decision-making. J. Theor. Biol. 189, 211–225 (1997). https://doi.org/10.1006/jtbi.1997.0511
[52] Dawson, D. A. Introductory lectures on stochastic population systems. arXiv (2017). https://doi.org/10.48550/arXiv.1705.03781
[53] Felsenstein, J. Macroevolution in a model ecosystem. Am. Nat. 112, 177–195 (1978). https://doi.org/10.1086/283259
[54] Nee, S. Birth-death models in macroevolution. Annu. Rev. Ecol. Evol. Syst. 37, 1–17 (2006). https://doi.org/10.1146/annurev.ecolsys.37.091305.110035
[55] Pennell, M. W., Harmon, L. J., Uyeda, J. C. Is there room for punctuated equilibrium in macroevolution? Trends Ecol. Evol. 29, 23–32 (2014). https://doi.org/10.1016/j.tree.2013.07.004