The application of statistical physics to evolutionary biology

Guy Sella[†‡] and Aaron E. Hirsh[§]

[†]Center for the Study of Rationality, Hebrew University of Jerusalem, Givat Ram, Jerusalem 91904, Israel; and [§]Department of Biological Sciences, Stanford University, Stanford, CA 94305

Edited by Simon A. Levin, Princeton University, Princeton, NJ, and approved May 7, 2005 (received for review March 7, 2005)

A number of fundamental mathematical models of the evolutionary process exhibit dynamics that can be difficult to understand analytically. Here we show that a precise mathematical analogy can be drawn between certain evolutionary and thermodynamic systems, allowing application of the powerful machinery of statistical physics to analysis of a family of evolutionary models. Analytical results that follow directly from this approach include the steady-state distribution of fixed genotypes and the load in finite populations. The analogy with statistical physics also reveals that, contrary to a basic tenet of the nearly neutral theory of molecular evolution, the frequencies of adaptive and deleterious substitutions at steady state are equal. Finally, just as the free energy function quantitatively characterizes the balance between energy and entropy, a free fitness function provides an analytical expression for the balance between natural selection and stochastic drift.

genetic drift | genetic load | neutral theory | steady state | fundamental theorem of natural selection

Even very simple mathematical models of the evolutionary process can be surprisingly difficult to understand analytically. For example, the Wright–Fisher process with viability selection, a relatively basic set of rules modeling stochastic reproduction and selection, gives rise to a probability distribution of allele frequencies that was revealed only through Kimura's application of diffusion theory and his solution of the resulting set of differential equations (1, 2). Similarly, Fisher's well known geometric model of evolution (3), in which adaptive change is represented by stepwise movement of a point toward the center of a hypersphere, has been studied by moment approximation and simulation (4–6), but exact analytical expressions for many quantities of interest remain inaccessible. In view of the analytic difficulties presented by our most basic models of evolution, new approaches that render large families of models more accessible could prove important both in furthering our understanding of the evolutionary process and in producing basic theoretical results useful in population genetic analysis of sequence data. Here we show that statistical physics furnishes one such approach.

Historical efforts to apply the methods of physics to the problems of biology can be divided into two rather different pursuits. In one, organisms, populations, or ecosystems are viewed as systems that, despite their emergent complexity, are subject to physical laws operating at lower levels of organization (7, 8). Attempts are then made to move from a lower level of organization, at which a physical system is analyzed, to higher levels of organization, at which biological systems are observed. This is not the approach we adopt here. In a very different pursuit, a parallel is drawn between a well understood physical system and a reduced or abstracted biological system; if it is sufficiently complete, the parallel allows the application of tools developed in the physical sciences to the analysis of biological systems. Successful examples of this approach include Kimura's application of diffusion theory (2) and Hopfield's analogy between neural networks and spin glasses (9, 10). Here we show that a very precise mathematical analogy can be developed between certain evolutionary and thermodynamic systems. This is a useful finding, because it allows us to apply the powerful tools of statistical physics to the analysis of simple evolutionary models, yielding several results.

In the present work, we concentrate on the family of models that depict the evolutionary process as a succession of mutant fixations, each of which occurs on the genetic background of the population's previous common ancestor. These models neglect linked polymorphism and the possibility of temporally overlapping fixations. Such effects are treated in other families of population genetic models (e.g., refs. 11–13), but we reserve for future work the extension of the methods developed here to those important problems. The successive fixation models examined here provide a decent approximation to the realistic population dynamic in systems in which the fixation probability of a mutation is not affected by other segregating alleles, with the obvious exception of the allele from which the new mutant was derived. (This condition holds when the product of the population size and the mutation rate is small, i.e., $N\mu \ll 1$.) Perhaps more importantly, as we will consider in the *Discussion*, the models examined here provide natural null models for nearly neutral evolution, with which alternative models involving more complex processes can be compared.

A general result that derives from the application of statistical physical methods to simple evolutionary models contradicts a basic tenet of the nearly neutral theory of molecular evolution. We therefore briefly review here the history and significance of that basic assumption. If the majority of evolutionary substitutions are truly neutral, the molecular divergence between two species is expected to be proportional to the number of generations that have elapsed since their separation (14). This prediction is contradicted by the observation that the rate of evolution appears to be roughly constant across organisms with dramatically different generation times (ref. 15, p. 38). To explain this relatively constant rate of evolution (among several other observations), Ohta (16) suggested that ". . . the majority of the amino acid substitutions in evolution, although subject to random genetic drift, are not completely neutral but rather very slightly selected against." If organisms with shorter generation times also have larger populations, Ohta reasoned, the reduced probability of fixation of slightly deleterious mutations in larger populations could offset the larger number of generations per year, resulting in a rate of evolution that does not depend on generation time.

It is important to note that Ohta's (16) suggestion that most substitutions are slightly deleterious does not lead necessarily to the (rather absurd) notion that all organisms are experiencing an ineluctable decline from an original state of perfect adaptation. The alternative to such steady decay is simply that each adaptive fixation compensates for many slightly deleterious fixations,

This paper was submitted directly (Track II) to the PNAS office.

[‡]To whom correspondence should be addressed. E-mail: gsella@math.huji.ac.il.

© 2005 by The National Academy of Sciences of the USA

resulting in a long-term steady state in which most substitutions are slightly deleterious (ref. 6; ref. 15, p. 38; and ref. 16). In effect, many baby steps downhill in the fitness landscape are offset by a larger leap uphill. This scenario is compatible with Ohta's (16) demonstration that in a reasonably large population, fixation of slightly deleterious mutations is possible, but fixation of more significantly deleterious mutations is not.

Models and Analysis

When a mutant appears in a population, its lineage faces two possible fates. The lineage may grow to take over the population, such that the mutant becomes the population's most recent common ancestor, or, as happens more frequently, the lineage may go extinct. When a single mutant with genotype j appears in a population with a wild-type genome i, the probability that it fixes depends on its fitness f_j, the fitness of the wild-type f_i, and the effective population size N. The probability of fixation also depends on the sampling process that describes the replacement of parents by offspring. For the Moran birth–death process, the exact probability of fixation is

$$\pi(i \rightarrow j) = \frac{1 - \frac{f_i}{f_j}}{1 - \left(\frac{f_i}{f_j}\right)^N} \qquad [1]$$

(see ref. 17, equation 3.50). In *Supporting Text*, which is published as supporting information on the PNAS web site, we show that for the Wright–Fisher process, the equation

$$\pi(i \rightarrow j) = \frac{1 - \left(\frac{f_i}{f_j}\right)^a}{1 - \left(\frac{f_i}{f_j}\right)^{2N}}, \qquad [2]$$

where $a = 2$ in a haploid population, and $a = 1$ for a diploid population with multiplicative fitness within loci, provides a closer approximation to the probability of fixation than the canonical formula (18). This expression also has properties that will prove convenient in analyses presented below.

Given the probability of fixation, the succession of mutant fixations can be depicted as a Markov process. In the models we study, the state of the population is described by the fixed genotype or, equivalently, by the genotype of the most recent common ancestor. The state of the Markov system at time t is described by a probability vector $\vec{P}(t) = (P_1(t), \ldots, P_M(t))$, where $P_i(t)$ denotes the probability that the most recent common ancestor of the population at time t had genotype i. The transition matrix W, describing the rate at which genotype j replaces genotype i as the most recent common ancestor of the population, is given by

$$W_{j,i} = \begin{cases} \frac{2}{a} C_{i,j} \mu_{i,j} N \pi(i \rightarrow j) & i \neq j \\ 1 - \sum_{k \neq i} W_{k,i} & i = j \end{cases}, \qquad [3]$$

where $C_{i,j}$ is 1 if genotype j is one mutation away from genotype i and is otherwise 0; $\mu_{i,j}$ is the rate of mutation from genotype i to genotype j; and a, defined above, depends only on the population's ploidity. Thus, the Markovian evolutionary dynamic in discrete time with discrete states takes the form $\vec{P}(t+1) = \mathbf{W}\vec{P}(t)$. Although here we use the Markov process that is discrete in both states and time, in *Supporting Text*, we show that our derivations apply also to the corresponding Markov systems that are continuous in states or time.

Table 1. The detailed analogy between evolutionary dynamics and statistical physics

Object	Evolutionary dynamic	Statistical physics
State variable	$\vec{g} = (A, T, T, \ldots, G)$	$\vec{s} = \{(\vec{q}_k, \vec{p}_k)\}$
Additive fitness and energy	$x = \ln(f(\vec{g}))$	$E = \hat{H}(\vec{s})$
Population size and temperature	$\nu_{\text{Moran}} = N - 1$ $\nu^h_{WF} = 2(N-1)$ $\nu^d_{WF} = 2N - 1$	$\beta = 1/k_B T$
Boltzmann factor	$P^i_{s.s.} \propto e^{-\nu(-x_i)}$	$P^i_{eq} \propto e^{-\beta E_i}$
Invariance	$f_i \rightarrow C f_i$	$E_i \rightarrow E_i + C$
Free fitness and free energy	$G = \langle x \rangle + \frac{1}{\nu} H$	$-G = -\left(\langle E \rangle - \frac{1}{\beta} H\right)$
Equilibrium scale	$\nu(x_j - x_i) = 1$	$\beta(E_j - E_i) = 1$

State variable, Additive fitness and energy, Population size and temperature, Boltzmann factor, and Free fitness and free energy are explained in the text. (Invariance) The analogy is also reflected in the symmetries in the representation of physical and evolutionary systems. Namely, because the representation of a physical system is invariant to the addition of a constant to the energy of all the microscopic states, the evolutionary system is invariant to multiplying the fitness in the system by a constant. The invariance takes precisely the same form if we replace fitness by the additive fitness, which is analogous to energy.

As a first step in developing the analogy between evolutionary and thermodynamic systems, we examine their equilibria. The sites in a gene that can be in one of four states (A, G, C, or T) are analogous to the degrees of freedom in statistical physics, such as the positions and momenta of particles in a gas (Table 1, State variable). The Markov system we have described is irreducible, i.e., there is a finite path of nonvanishing probability between any two states, or genotypes, in the system. This implies [by the Perron–Frobenius theorem (19)] that whatever the initial state, the distribution of probability that the population's most recent common ancestor is of any given genotype approaches a unique steady state. This stationary distribution is defined by the requirement that $\vec{P}^* = W\vec{P}^*$, or equivalently, that the probability flow in and out of each state is balanced, i.e.,

$$\sum_j (W_{j,i} P^*_i - W_{i,j} P^*_j) = 0, \qquad [4]$$

for any genotype i.

Because a number of distinct notions of the steady-state distribution have been used in the population genetic literature (1, 4, 20), we pause to clarify both the meaning of the steady state defined by Eq. **4** and its relation to the classical notion of mutation-selection-drift balance. A well known example of mutation-selection-drift balance is Kimura's U-shaped distribution, which describes the frequency distribution of an allele in a model with two neutral alleles that mutate into each other (1). When we wish to generalize the notion of Kimura's steady state to a system with many different genotypes at various mutational distances from one another, it is useful to view the U-shaped distribution as a superposition of two processes. The first process is that of fixation, which in the long run switches between the two alleles. Although the population may never be homogeneous at a locus, fixation is mathematically and conceptually well defined in terms of the state of the population's most recent common ancestor. The second process describes the frequency of the mutant allele conditional on the wild type being fixed. When both processes are combined, we obtain the U-shaped distribution. However, when we consider a system with many genotypes at various mutational distances from one another and exhibiting

various selection coefficients, calculating the generalization of the U-shaped distribution becomes very difficult. Moreover, such a generalization may not be very interesting, simply because very few of the alleles that appear in such a distribution appear together in the population at the same time (in general, only alleles that are one, or very few, mutations from each other appear in the population simultaneously). A natural alternative to this exceedingly complex steady-state distribution is the distribution of time spent with each allele fixed or, equivalently, the probability distribution of finding the population with a given allele fixed. In principle, any population genetic measure of interest, such as the average fitness (see below), average heterozygosity, or the effective number of alleles, can be calculated conditional on this steady-state distribution. For example, in the limit where no more than two alleles exist simultaneously in the population ($N\mu \ll 1$), the sojourn time formalism (17) can be used to calculate heterozygosity conditional on any given fixed genotype, and the steady-state distribution can be used to calculate the average heterozygosity across the fixed genotypes. This form of steady state has recently been used in the literature (4, 6), and it is precisely the steady state we calculate here. In the case of Kimura's system with two neutral alleles, this steady state is simply (1/2, 1/2). Naturally, the distribution becomes substantially more complicated when one considers multiple genotypes at various mutational distances from one another and exhibiting various selection coefficients.

Quite surprisingly, the steady-state distribution of fixed genotypes can be found for any given fitness scheme. Here we assume that mutation is symmetric, i.e., that $\mu_{i,j} = \mu_{j,i}$; in *Supporting Text*, we generalize our results to cover cases of asymmetric mutation. Consider a single edge in the fitness landscape, specifically, the edge connecting genotypes i and j, which are separated by a single mutation. From Eqs. **1–3**, we see that for this edge,

$$\frac{W_{j,i}}{W_{i,j}} = \frac{(f_j)^v}{(f_i)^v} = \frac{F(j)}{F(i)}, \qquad [5]$$

where $v = N - 1$ for the Moran process, $v = 2(N - 1)$ for the haploid Wright–Fisher process, and $v = 2N - 1$ for the diploid Wright–Fisher process with multiplicative fitness within a locus. When the rates of transition between any two states satisfy such a relation, in which their ratio can be described as a ratio of a function F evaluated at the two states, statistical physics tells us that the steady-state distribution takes a simple form (19). Because Eq. **5** shows that

$$W_{j,i}\frac{F(i)}{\sum_K F(k)} = W_{i,j}\frac{F(j)}{\sum_K F(k)}, \qquad [6]$$

it follows that the unique solution to Eq. **4** is given by

$$P_i^* = \frac{F(i)}{\sum_K F(k)} = \frac{(f_i)^v}{\sum_K (f_k)^v} = \frac{e^{vx_i}}{\sum_K e^{vx_k}}, \qquad [7]$$

where $x_I = \ln(f_i)$. We introduce the variable x, which we refer to as the additive fitness, for two reasons. First, the additive fitness x, unlike f, exhibits a desirable property when we focus on the relationships among multiple alleles. As a simple example, consider three alleles that have fitness values f_1, f_2, and f_3. The usual population genetic selection coefficient of allele 2 relative to allele 1 would be $s_{1,2} = f_2/f_1 - 1$. Similarly, the selection coefficient of allele 3 relative to allele 2 would be $s_{2,3} = f_3/f_2 - 1$. Consideration of allele 3 relative to allele 1 shows that the selection coefficient, s, does not behave additively. That is, $s_{1,3} \neq s_{1,2} + s_{2,3}$. However, if we use x instead of f, the relationships are simplified. Instead of the usual population genetic selection coefficient $s_{i,j} = f_j/f_i - 1$, we now have $\Delta x_{i,j} = x_j - x_i$ and, conveniently, $\Delta x_{1,3} = \Delta x_{1,2} + \Delta x_{2,3}$. Moreover, this additivity extends to an arbitrary number of alleles.

The second reason for introducing the variable x is that the substitution places the steady-state distribution (Eq. **7**) in a form familiar from statistical physics, emphasizing that this steady state is precisely analogous to the description of a physical system at thermal equilibrium. The analogy is summarized in Table 1. When a physical system is at equilibrium with a thermal bath at temperature T, the probability that the system would be in a microscopic state i is given by the Boltzmann factor,

$$P_i^* \propto e^{-\beta E_i}, \qquad [8]$$

where $\beta = 1/k_B T$, k_B is the Boltzmann coefficient, and E_i is the energy of state i (21). Similarly, we find that the probability of the population being fixed with genotype i at steady state is given by

$$P_i^* \propto e^{-v(-x_i)}, \qquad [9]$$

where v is analogous to β, and the additive fitness, $-x_i$, is analogous to the energy E_i. Energy is an additive quantity in physical systems, so it is not surprising that its counterpart in the evolutionary process should be the additive fitness, x. (Because the sign of energy is defined such that the dynamics tend to reduce it, whereas evolutionary systems tend to increase fitness, a minus sign must be introduced in translation.) The analogy between v and β indicates that the population size affects the evolutionary system as the inverse of the temperature affects a physical system. When the temperature is zero, a physical system at equilibrium is always at the lowest energy state. Analogously, when the population is infinite, the evolutionary system at steady state is always at the genotype with the highest fitness. When the temperature is not zero, the probability of finding the physical system at any microscopic state depends only on a state's energy. Analogously, in an evolutionary system with a finite population, the probability of finding the population with a given fixed genotype depends only on the fitness associated with that genotype.

Before developing the analogy further, we pause to consider its use in addressing a few basic questions about the nature of the steady-state distribution of fixed genotypes. At this steady state, the evolutionary system is constantly changing, but without adaptation (6, 22). As discussed in the Introduction, a basic tenet of nearly neutral theory is that many small deleterious substitutions are compensated by fewer more considerable adaptations. This turns out to be incorrect. An important property that derives from the steady-state solution is that along any edge in the landscape

$$W_{j,i}P_i^* = W_{i,j}P_j^*. \qquad [10]$$

This property, which plays a central role in statistical physics, is called detailed balance (DB) (19). For any system at steady state, the flow of probability in and out of each state is balanced. For a system that satisfies DB, this balance holds between any two states. Thus, at steady state, the rate at which one genotype fixes and replaces the other as wild type is precisely equal to the rate at which the opposite fixation occurs; this implies that the numbers of adaptive and deleterious substitutions in the evolutionary system are equal. The implications of this equality for studies of nearly neutral evolution will be considered in the *Discussion*.

Because the steady-state distribution of fixed genotypes includes types that are not maximally fit, the population suffers a cost. This cost is referred to as the fixed-drift load (4), and it is generally calculated as the proportional reduction of the population's average fitness due to stochastic fixation of suboptimal

alleles. Calculation of load bears on important population genetic problems, such as the evolution of sex and the extinction risk of small populations (4, 23, 24). The analogy developed here reduces the calculation of fixed-drift load to several straightforward steps, into which appropriate expressions may be substituted according to the particular population genetic model or fitness landscape under consideration. For brevity, we illustrate these steps using a simple model, corresponding to Fisher's geometric model in one dimension; application of the same approach to complex models, such as Fisher's multidimensional model, as well as a detailed comparison of the analytic results with simulations, will be demonstrated elsewhere. Consider the simple landscape in which fitness ranges continuously between 0 and 1, and the density of genotypes that correspond to any fitness is uniform, i.e., $\rho_g(f) = 1$ for $f \in (0,1)$. Assume further that mutation between genotypes is symmetric. According to Eq. **7**, at steady state, the probability density of having fitness f is

$$\rho^*(f) = \frac{f^v \rho_g(f)}{\int_0^1 f^v \rho_g(f) df}. \qquad [11]$$

Therefore, the average fitness is

$$\langle f^* \rangle = \frac{\int_0^1 f f^v \rho_g(f) df}{\int_0^1 f^v \rho_g(f) df} = \frac{v+1}{v+2}, \qquad [12]$$

and the fixed-drift load is

$$L = \frac{f_{\max} - \langle f^* \rangle}{f_{\max}} = \frac{1}{v+2}. \qquad [13]$$

That the fixed-drift load is approximately $1/v$ becomes intuitive below, when we consider how the efficacy of selection at steady state depends on the effective population size. Eq. **13** shows that the load is independent of the properties of mutation other than symmetry; this follows from the more general result that the steady-state distribution of fixed genotypes (Eq. **7**) is itself independent of the properties of mutation. This distribution (and, consequently, quantities that can be expressed as functions of this distribution alone) depend only on the fitness function and the population size. In *Supporting Text*, we show that these results still hold when we incorporate the well established asymmetries in mutation, such as those responsible for GC or AT bias (25). The independence of steady state from the details of mutation considerably simplifies the analysis of evolutionary behavior at equilibrium.

What can we say about the evolutionary process before steady state is attained? In physics, it is useful to find an energy function (or a Lyapunov function in mathematics), i.e., a function of the system's state that changes monotonically as the dynamics progress in time. In *Supporting Text*, we use detailed balance to show that

$$G = \langle \ln(f) \rangle + \frac{1}{v} S, \qquad [14]$$

where $S = -\langle \ln(P) \rangle$, is an energy, or Lyapunov, function of the evolutionary dynamic. [The proof that G monotonically increases under the evolutionary dynamics is precisely analogous to the proof of Boltzmann's H theorem (21).] Because of its close parallel with free energy, we refer to Eq. **14** as the free fitness function (Table 1, Free fitness, free energy). The monotone increase of free fitness and its eventual maximization at steady state is analogous to the maximization of free energy at thermal equilibrium, which is a manifestation of the second law of thermodynamics. At thermal equilibrium, the minimization of free energy balances between a physical system's tendencies to lower energy and increase entropy (S). At the evolutionary steady state, the maximization of free fitness balances between the evolutionary tendencies in finite populations to increase both fitness and entropy. Although other Lypunov functions can be found, this function is uniquely defined by being extensive (26, 27), i.e., if we consider a system of n genetic sites that contribute to fitness independently, the free fitness of this system equals the sum of free fitness values across sites.

The monotone increase of free fitness bears an interesting relationship to Fisher's fundamental theorem of natural selection (3), which states that a population's average fitness increases at a rate proportional to the additive genetic variance in the population. [To be precise, Fisher's theorem has been interpreted (28–30) as stating that in an infinite population, the fitness component associated with a constant environment, both ecologically and genetically, increases at a rate proportional to the additive genetic variance.] Unlike the models considered here, Fisher's theorem applies under conditions of a changing environment and includes an explicit representation of a population's standing genetic variance. Despite its extraordinary generality, however, Fisher's theorem does make the important simplifying assumption that the population size is infinite; it omits the effects of drift in finite populations. The free fitness function, although it lacks the generality of Fisher's theorem on other counts, does account for the evolutionary effects of finite population size. When the population size is infinite, the second term in the free fitness function vanishes, and the function's monotone increase becomes an increase in average fitness, in accordance with Fisher's theorem. But when the population is finite, the second term in the free fitness function is nonvanishing. Indeed, this term, which is simply the entropy of the probability distribution of genotypes, becomes proportionally more important as the population size becomes smaller. Whereas increases in the first term of the free fitness function are associated with adaptation, increases in the second term indicate augmented importance of stochastic, or nonadaptive, change.

Fisher (ref. 3, p. 39), as well as others after him (e.g., ref. 6), has wondered about the relative efficacy of the forces "destroying adaptation" and that "building it up." To put the tradeoff between fitness and entropy in quantitative terms, we consider how population size affects selective discrimination between genotypes at steady state. The scale of selective discrimination, i.e., the difference in fitness that causes a unit of change in the logarithm of the probability of a genotype at steady state, derives from the relation

$$\ln\left(\frac{P_i^*}{P_j^*}\right) = v \ln\left(\frac{f_i}{f_j}\right) = 1. \qquad [15]$$

For a biologically realistic population size ($N \gg 1$), Eq. **15** implies that

$$\ln\left(\frac{f_i}{f_j}\right) = \ln(1 + s_{j,i}) \approx s_{j,i}, \qquad [16]$$

where $s_{j,i}$ is the selection coefficient. Therefore, the scale of selective discrimination satisfies the relation $vs \approx 1$. That is, free fitness determines the tradeoff between fitness and entropy such that a selection coefficient of $1/v$ causes a unit change in the steady-state probability of a genotype. This relation, which is a generalization of classical results from neutral theory (14, 31), distills how population size determines the efficacy of selection.

This is analogous to a thermodynamic system, in which the temperature determines how differences in the energy associated with different states affect their probabilities at thermal equilibrium (Table 1, Equilibrium scale).

Discussion

The suggestion that most amino acid substitutions are slightly deleterious was originally offered to resolve contradictions between the predictions of neutral theory and empirical observations (15). In exploring how this novel premise might realign predictions and data, a number of theoretical studies simulated molecular evolution under predefined distributions of mutant selection coefficients (see, for example, ref. 32). The results presented here suggest that *a priori* assumption of a particular distribution of mutant selection coefficients is inappropriate for steady-state models of nearly neutral evolution. The distribution of mutant selection coefficients is determined by the operation of the evolutionary dynamic on the landscape of all possible genotypes and therefore cannot be assumed *a priori*. In fact, the distribution determined by the evolutionary dynamic will differ in important ways from distributions assumed *a priori*. For instance, we found that the steady-state distribution of mutant selection coefficients will take a form that causes the number of fixations involving an additive fitness change Δx to equal the number of fixations involving an additive fitness change $-\Delta x$. This property will not, in general, be exhibited by distributions of mutant selection coefficients assumed *a priori*.

Although we find that the premise that most substitutions are slightly deleterious may be incorrect, many of the deductions that rely on this premise remain in tact. For example, steady-state nearly neutral evolution may still provide an explanation for the constancy of the rate of molecular evolution across organisms with very different generation times. According to detailed balance, on average, for one adaptive substitution with a given selection coefficient to occur, a corresponding deleterious substitution must also take place. Because deleterious mutations are much less likely to fix than adaptive ones, fixation of deleterious mutations remains the rate-limiting step in molecular evolution, and Ohta's (16) rationale for the molecular clock may still hold. Other classical arguments, such as Ohta's explanation for the surprisingly weak relationship between polymorphism and population size, can be rephrased along similar lines.

An important simplifying assumption in our analysis was that mutations are sufficiently rare that the fixation probability of a mutation is not affected by other segregating alleles. Although future work may determine whether the methods presented here can be generalized to incorporate interactions among segregating sites, we must presently address the applicability of theory that omits such important processes as hitchhiking (15, 33) and background selection (12). Most straightforwardly, the results presented here can be viewed as an approximation that becomes acceptably accurate under certain population parameters. For example, in small populations, mutations are rare, and therefore our simplified evolutionary dynamics and the results we have derived from them become decent approximations of reality (4). More generally, however, the theory considered here may allow us to more adequately understand the behavior of models that frequently serve either as null models or as theoretically tractable heuristics in studies of more complicated molecular evolutionary dynamics. For instance, revising the premise that slightly deleterious substitutions predominate may affect how the null hypothesis of steady-state nearly neutral evolution is simulated in studies measuring rates and effects of adaptation in the genome. Or, to offer another example, closed-form expressions for the effect of population size on key population genetic statistics may facilitate evaluation of the suggestion that the effects of hitchhiking or background selection can be approximated theoretically by a rescaling of population size in neutral theory (12, 34).

Besides adopting important simplifying assumptions, the theory treated here reduces the complexity of dynamics by treating the averages of large numbers of degrees of freedom. Although such averages are clearly meaningful in the classical systems of statistical physics, in which the number of degrees of freedom is generally on the order of Avogadro's number, one may reasonably ask whether averages are actually useful in genetic systems, in which the number of degrees of freedom is so much smaller. Although population genetics initially concerned itself with the dynamics at one or a few sites of interest, the recent flood of genomic data has allowed measurement of averages over many sites. For instance, it is now common to compare average evolutionary rates or average levels of codon bias across many genes in the genome (35, 36). Such studies encounter tremendous noise, and predictions are far from precise (especially by the standards of statistical physics), but averaging across sites within genes and comparing large numbers of genes often allow detection of important trends.

The applicability of the theory treated here should also be discussed in light of Eq. **9**, which gives the probability that any given genotype is fixed in the population. The time scales of evolution do not allow exhaustive exploration of large and rugged fitness landscapes; instead, populations are often confined to exploration of a local fitness peak. Such metastable states do not reflect an equilibrium in the strict sense because, given sufficient time, the system would eventually leave each local peak. Nonetheless, under the approximation that the local landscape is bounded by valleys that cannot be traversed, Eq. **9** and the results that follow from it would apply.

A number of authors (3, 37–39) have noted parallels between statistical physics and evolution. In 1930, R. A. Fisher wrote that ". . . the fundamental theorem [of natural selection] bears some remarkable resemblance to the second law of thermodynamics. It is possible that both may ultimately be absorbed by some more general principle" (3). In accordance with this suggestion, we have shown that the mathematical description of evolution of a finite population in a constant environment is analogous to that of a thermodynamic system. Our treatment does not address how evolutionary systems, which are themselves physical systems, are subject to the laws of statistical physics [as, for example, Schrödinger suggests (40)]. The analogy we have developed does, however, show that the methods used to describe systems in statistical physics can be applied to evolutionary systems in a useful way. This analogy leads to a general analytical form for the steady-state distribution of fixed genotypes, thus reducing the solution of a large family of evolutionary models, including Fisher's geometric model, to several straightforward substitutions. The close parallels between statistical physics and evolutionary dynamics also prove useful in elucidating basic evolutionary relationships, such as that between genetic load and effective population size, and in revealing new generalizations about dynamic behavior at steady state, such as the equality of the number of adaptive and deleterious substitutions. Finally, the analogy permits derivation of an energy function for evolutionary dynamics of a finite population. The form of this energy function is precisely that of free energy, and the maximization of free fitness is precisely analogous to the second law of thermodynamics.

We thank Ilan Eshel for numerous helpful discussions. We also thank Tsvi Tlusty, Nati Linial, Erel Levin, Itamar Pitowsky, Sergiu Hart, Yossi Rinott, Dmitri Petrov, and two anonymous reviewers for valuable comments at various stages of this work. This work was supported by the Center for Complexity Science of the Yeshaya Horowitz Association (G.S.). A.E.H. was supported by National Institutes of Health Grant 28428 (to Marcus W. Feldman).

1. Kimura, M. (1955) *Proc. Natl. Acad. Sci. USA* **41,** 144–150.
2. Kimura, M. (1955) *Cold Spring Harbor Symp. Quant. Biol.* **20,** 33–53.
3. Fisher, R. A. (1958) *The Genetical Theory of Natural Selection* (Dover, New York).
4. Poon, A. & Otto, S. P. (2000) *Evolution (Lawrence, Kans.)* **54,** 1467–1479.
5. Orr, H. A. (1999) *Genet. Res. Camb.* **74,** 207–214.
6. Hartl, D. L. & Taubes, C. H. (1996) *J. Theor. Biol.* **182,** 303–309.
7. Brown, J. H., Gupta, V. K., Li, B. L., Milne, B. T., Restrepo, C. & West, G. B. (2002) *Philos. Trans. R. Soc. London B* **357,** 619–626.
8. West, G. B., Woodruff, W. H. & Brown, J. H. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 2473–2478.
9. Hopfield, J. J. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 2554–2558.
10. Hertz, J., Krogh, A. & Palmer, R. G. (1991) *Introduction to the Theory of Neural Computation* (Addison–Wesley, New York).
11. Braverman, J. M., Hudson, R. R., Kaplan, N. L., Langley, C. H. & Stephan, W. (1995) *Genetics* **140,** 783–796.
12. Charlesworth, B., Morgan, M. T. & Charlesworth, D. (1993) *Genetics* **134,** 1289–1303.
13. Zhivotovsky, L. A., Feldman, M. W. & Christiansen, F. B. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 1079–1083.
14. Kimura, M. & Ohta, T. (1971) *Nature* **229,** 467–469.
15. Gillespie, J. H. (2004) *Population Genetics* (Johns Hopkins Univ. Press, London), 2nd Ed.
16. Ohta, T. (1972) *J. Mol. Evol.* **1,** 150–157.
17. Ewens, W. J. (1979) in *Mathematical Population Genetics*, eds. Krickeberg, K. & Levin, S. A. (Springer, Berlin).
18. Crow, J. F. & Kimura, M. (1970) *An Introduction to Population Genetics Theory* (Harper & Row, New York).
19. van Kampen, N. G. (1981) *Stochastic Processes in Physics and Chemistry* (North–Holland, Amsterdam).
20. Wright, S. G. (1945) *Proc. Natl. Acad. Sci. USA* **24,** 372–377.
21. Reichl, L. E. (1980) *A Modern Course in Statistical Physics* (University of Texas Press, Austin).
22. Ohta, T. (1973) *Nature* **246,** 96–98.
23. Muller, H. J. (1964) *Mutat. Res.* **1,** 2–9.
24. Lynch M., Bürger, R., Butcher, D. & Gabriel, W. (1993) *J. Hered.* **84,** 339–344.
25. Bernardi, G. (1986) *J. Mol. Evol.* **24,** 1–11.
26. Kats, A. (1967) *Principles of Statistical Mechanics* (Freeman, San Francisco).
27. Shannon, C. E. & Weaver, W. (1963) *The Mathematical Theory of Communication* (University of Illinois Press, Urbana, IL).
28. Price, G. R. (1972) *Ann. Hum. Genet.* **36,** 129–140.
29. Ewens, W. J. (1989) *Theor. Popul. Biol.* **36,** 167–180.
30. Frank, S. A. & Slatkin, M. (1992) *TREE* **7,** 92–95.
31. Kimura, M. (1983) *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge, U.K.).
32. Ohta, T. & Tachida, H. (1990) *Genetics* **126,** 219–229.
33. Maynard Smith, J. & Haigh, J. (1974) *Genet. Res. Camb.* **23,** 23–35.
34. Gillespie, J. H. (2001) *Evolution (Lawrence, Kans.)* **55,** 2161–2169.
35. Wall, D. P., Hirsh, A. E., Fraser, H. B., Kumm, J., Gianer, G., Eisen, M. B. & Feldman, M. W. (2005) *Proc. Natl. Acad. Sci. USA* **102,** 5483–5488.
36. Powell, J. R. & Moriyama, E. N. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 7784–7790.
37. Kauffman, S. A. & Weinberger, E. D. (1989) *J. Theor. Biol.* **141,** 211–245.
38. Derrida, B. & Peliti, L. (1991) *Bull. Math. Biol.* **53,** 355–382.
39. van Nimwegen, E., Crutchfield, J. P. & Huynen, M. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 9716–9720.
40. Schrödinger, E. (1967) *What Is Life?* (Cambridge Univ. Press, Cambridge, U.K.).

Corrections

BIOCHEMISTRY. For the article "Kernel energy method: Application to insulin," by Lulu Huang, Lou Massa, and Jerome Karle, which appeared in issue 36, September 6, 2005, of *Proc. Natl. Acad. Sci. USA* (**102,** 12690–12693; first published August 24, 2005; 10.1073/pnas.0506378102), the authors note an error in a funding acknowledgment: "L.M. thanks the National Institutes of Health for National Institute of General Medical Sciences Grant MBRS SCORE5 S06GM606654 and National Center For Research Resources Grant RR-0307 and the National Science Foundation for Centers of Research Excellence in Science and Technology grant support" should read: "L.M. thanks the National Institutes of Health for National Institute of General Medical Sciences Grant MBRS SCORE5 S06GM606654 and the National Science Foundation for Centers of Research Excellence in Science and Technology grant support. This investigation was supported by Research Centers in Minority Institutions' Award RR-03037 from the National Center for Research Resources, National Institutes of Health."

www.pnas.org/cgi/doi/10.1073/pnas.0507559102

BIOPHYSICS. For the article "Random-coil behavior and the dimensions of chemically unfolded proteins," by Jonathan E. Kohn, Ian S. Millett, Jaby Jacob, Bojan Zagrovic, Thomas M. Dillon, Nikolina Cingel, Robin S. Dothager, Soenke Seifert, P. Thiyagarajan, Tobin R. Sosnick, M. Zahid Hasan, Vijay S. Pande, Ingo Ruczinski, Sebastian Doniach, and Kevin W. Plaxco, which appeared in issue 34, August 24, 2004, of *Proc. Natl. Acad. Sci. USA* (**101,** 12491–12496; first published August 16, 2004; 10.1073/pnas.0403643101), the authors note the following: "Due to a mathematical error in our analysis, we incorrectly reported the prefactor (R_0) for the Flory scaling relationship observed between the experimental R_G of chemically denatured proteins and their sequence length:

$$R_G = R_0 N^{\nu}.$$

The correct, best-fit value for the prefactor R_0 is $1.927^{+0.271}_{-0.238}$ Å (bounds represent 95% confidence intervals). We apologize for any inconvenience this error may have caused. This error does not affect the conclusions of the article."

www.pnas.org/cgi/doi/10.1073/pnas.0507472102

EVOLUTION. For the article "The application of statistical physics to evolutionary biology," by Guy Sella and Aaron E. Hirsh, which appeared in issue 27, July 5, 2005, of *Proc. Natl. Acad. Sci. USA* (**102,** 9541–9546; first published June 24, 2005; 10.1073/pnas.0501865102), the authors note the following: "The stationary distribution of fixed genotypes (Eq. **9**) was previously derived (1, 2) in the context of the evolution of transcription factor binding sites, by using detailed balance of the substitution dynamics (Eq. **3**). The formula appears slightly different between refs. 1 and 2 and ref. 3, because of slight differences in assumptions; in refs. 1 and 2, continuous time is assumed, whereas in ref. 3, discrete time is assumed. Furthermore, in refs. 1 and 2, fitness is parameterized in terms of molecular distance from an optimum sequence. These authors also discuss a linear combination of fitness and a quantity they call 'mutational entropy'; when the system is at equilibrium, this combination is very similar to our free fitness."

1. Berg, J. & Laessig, M. (2003) *Biophysics (Moscow)* **48,** Suppl. 1, S36–S44.
2. Berg, J., Willmann, S. & Laessig, M. (2004) *BMC Evol. Biol.* **4,** 42.
3. Sella, G. & Hirsh, A. E. (2005) *Proc. Natl. Acad. Sci. USA* **102,** 9541–9546.

www.pnas.org/cgi/doi/10.1073/pnas.0507361102

GENETICS. For the article "Loss and gain of chromosome 5 controls growth of *Candida albicans* on sorbose due to dispersed redundant negative regulators," by M. Anaul Kabir, Ausaf Ahmad, Jay R. Greenberg, Ying-Kai Wang, and Elena Rustchenko, which appeared in issue 34, August 23, 2005, of *Proc. Natl. Acad. Sci. USA* (**102,** 12147–12152; first published August 11, 2005; 10.1073/pnas.0505625102), on page 12152, the first sentence of the last paragraph in the left column, "If each region encompasses at least one *CSU* gene, *CSU51–CSU55*, then three more genes in regions B, C, and 139 are expected" should read: "If each region encompasses at least one *CSU* gene, *CSU51–CSU55*, then four more genes in regions B, 135, C, and 139 are expected." The conclusions of the article remain unchanged.

www.pnas.org/cgi/doi/10.1073/pnas.0507247102

Supporting Text

Derivation of the Probability of Fixation

In this supplement, we establish the validity of Eq. **2**. First, we show that the fixation probability in the Wright-Fisher process is well approximated by

$$\pi(i \to j) = \frac{1-(\frac{f_i}{f_j})^a}{1-(\frac{f_i}{f_j})^{2N}}, \qquad [\mathbf{17}]$$

where $a = 2$ in a haploid population, and $a = 1$ for a mutation with multiplicative fitness across loci in a diploid population. In fact, Eq. **17** provides a closer approximation than does the canonical fixation formula (1). Second, we explain the mathematical considerations that lead to Eq. **17**. Finally, we list some advantages of the new fixation formula.

Table 2 presents a comparison among probabilities of fixation determined in three different ways: (*i*) calculated according to the canonical fixation formula,

$$\pi(p, s, N) = \frac{1-e^{-2Nps}}{1-e^{-2Ns}}; \qquad [\mathbf{18}]$$

(*ii*) calculated according to the proposed formula (Eq. **15**), which can be written

$$\pi(p, s, N) = \frac{1-e^{-2Np\Delta x}}{1-e^{-2N\Delta x}}, \qquad [\mathbf{19}]$$

where $\Delta x = \ln(1+s)$; and (*iii*) measured in simulations of the Wright-Fisher process. The comparison was performed for a variety of population sizes, selective coefficients, and initial frequencies of the invading type. For each combination of parameters, evaluation of the probability of fixation was based on 10^7 simulated paths. Because the probability of fixation and therefore the accuracy of the evaluation decrease as a

function of the population size, we limit the comparison to small population sizes. It is apparent that for nearly neutral mutations, Eq. **19** provides a very accurate approximation to the measured probability of fixation, and that Eq. **19** is generally more accurate than the canonical formula (Eq. **18**).

The derivations of Eqs. **18** and **19** are very similar. The canonical fixation formula is derived as a diffusion approximation to the Wright-Fisher process (2). The derivation can be described in two steps. First, the following differential equation for the fixation probability is derived:

$$\frac{d^2\pi}{dp^2} + 2g(p,s,N)\frac{d\pi}{dp} = 0, \qquad [\mathbf{20}]$$

where

$$g(p,s,N) = \frac{E\{\Delta p | p\}}{E\{(\Delta p)^2 | p\}}. \qquad [\mathbf{21}]$$

This differential equation is an approximation that ignores all the moments in Δp above second order. For the haploid Wright-Fisher system, one can show that

$$g(p,s,N) = 2Ns\frac{1+sp}{2Ns^2 pq + (1+s)}, \qquad [\mathbf{22}]$$

where $q = 1 - p$. The second step in the derivation is to approximate the expression for g under the assumption that $s, \frac{1}{N}, Ns^2 << 1$, which yields

$$g(p,s,N) \approx 2Ns. \qquad [\mathbf{23}]$$

The canonical fixation formula (Eq. **18**) results from substituting expression **23** into differential equation (Eq. **20**) and incorporating the boundary conditions

$$\pi(0) = 0$$
$$\pi(1) = 1. \qquad [\mathbf{24}]$$

Substituting

$$g(p,\Delta x,N)\approx 2N\Delta x \qquad [\mathbf{25}]$$

into DE **20** yields fixation formula **19**. Note that Eqs. **18** and **19** are valid under the same conditions, namely, $s,\frac{1}{N},Ns^2<<1$, or equivalently, $\Delta x,\frac{1}{N},N\Delta x^2<<1$.

Several considerations motivated the use of the variable Δx rather than s. If we consider a composition of two or more mutations, which contribute to fitness independently, it would be mathematically desirable that the selective parameter for this composition would be exactly the sum of selective parameters for each of the mutations taken alone. Considering the simplest example of two independent mutations shows that Δx satisfies this requirement, whereas s does not. If s_1 and s_2 are the selective coefficients for two mutations, and, $\Delta x_1=\ln(1+s_1)$ and $\Delta x_2=\ln(1+s_2)$ are the corresponding selective factors, then the selective coefficient for the composition of these mutations is $\hat{s}=s_1+s_2+s_1s_2$, whereas the corresponding selective factor is $\Delta\hat{x}=\Delta x_1+\Delta x_2$. In other words, whereas Δx is an additive measure of selective advantage, s is not. Because x is an additive measure, when we consider the g that corresponds to a composition of mutations either as a function of the Δx s for each mutation or as a function of the Δx corresponding to the composition, we will always get the same result to any order in either of the Δx s. This will not be the case if we use s as the selection parameter. It is certainly not wrong to use s, because the fixation formulas are derived at the diffusion limit, at which second-order terms may be neglected, and therefore the parameter s may be considered additive to a good approximation. Nevertheless, as Table 1 shows, when a parameter is chosen for an expansion (in our case, the expansion of g), the approximation may be more reliable

when the parameter is chosen such that the natural operations of the system (in our case a composition of mutations) leave the orders of the expansion invariant.

To demonstrate how using the additive parameter for selection affects the properties of the expression for the probability of fixation, we consider the following consistency requirement. The fixation of an invading type with frequency p is equivalent to the extinction of the wild type with frequency $q = 1 - p$. It is therefore natural to require that an approximation of the probability of fixation should satisfy the condition that the probability of mutant fixation would always equal the probability of wild-type extinction. Because the additive fitness is additive, if the difference in additive fitness associated with an invading type is Δx, the difference in additive fitness associated with the wild type is $-\Delta x$. Eq. **19** then satisfies the consistency requirement

$$\hat{\pi}(p, \Delta x, N) = 1 - \hat{\pi}(q, -\Delta x, N). \quad [\mathbf{26}]$$

Correspondingly, if the selective coefficient of the invading type is s_{inv}, the selective coefficient for the wild type is

$$s_{wt} = -\frac{s_{inv}}{1 + s_{inv}}. \quad [\mathbf{27}]$$

The canonical fixation formula does not satisfy the consistency requirement because

$$\pi(p, s, N) \neq 1 - \pi(q, -\frac{s}{1+s}, N). \quad [\mathbf{28}]$$

As noted above, it is true that at the diffusion limit, $s_{wt} \approx -s_{inv}$, and therefore one can say that the canonical fixation formula satisfies this consistency requirement at the diffusion limit. However, it is both aesthetically pleasing and, as shown in Table 1, potentially useful to have an approximation of the fixation probability that satisfies the consistency requirement in the strict sense.

Another pleasing property of the fixation formula (Eq. **19**) for the Wright-Fisher process is that it has a form quite similar to the exact fixation formula for the Moran process. The factor 2 that distinguishes the formulas is accounted for simply by the 2-fold difference of the variance in the offspring distribution of the Moran process (2). But perhaps the most important advantage of the new fixation formula is that, in some circumstances, it allows for more analytical progress than the canonical one.

Extending the Derivations to Continuous Time and a Continuous Genotype Space

For simplicity, in the paper we have used Markov dynamics, which is discrete in both states and time. The generalization of our results to a continuous time and state is straightforward (3). In continuous time, the discrete dynamic $\vec{P}(t+1) = W\vec{P}(t)$ is replaced by

$$\frac{dP_i(t)}{dt} = \sum_k (W_{i,k}P_k(t) - W_{k,i}P_i(t)), \qquad [\mathbf{29}]$$

and with a continuous state space (genotypes in our case):

$$\frac{\partial P(i,t)}{\partial t} = \int (W(i \mid k)P(k,t) - W(k \mid i)P(i,t))dk, \qquad [\mathbf{30}]$$

where the transition matrix **W** is given by Eq. **3**, and when the state space is continuous, $\vec{P}$ becomes a vector of densities, and i and k become continuous parameterizations of the state space. A quick look at Eq. **6** shows that the steady-state solution we derive and detailed balance (DB) apply to both cases. The same is also true of the derivations of the free fitness function.

Proof That the Free Fitness Function Increases in Evolution

The proof that the free fitness function monotonically increases under the evolutionary dynamics is precisely analogous to the proof of Boltzmann's H theorem (7). To show that the free fitness function monotonically increases in the evolutionary dynamics, we simply show that its derivative according to time is non-negative. For that purpose, it is most convenient to use the formalism in continuous time. The derivative of the free fitness function according to time is

$$\begin{aligned}\frac{dG}{dt} &= \frac{d}{dt}\left(\sum_i P(i)\left(\ln(f_i) - \frac{1}{\nu}\ln(P(i))\right)\right) \\ &= -\frac{1}{\nu}\sum_i \frac{dP(i)}{dt}\left(\ln(P(i)) - \nu\ln(f_i) + 1\right) \\ &= -\frac{1}{\nu}\sum_i \frac{dP(i)}{dt}\left(\ln(\frac{P(i)}{P^*(i)}) + 1 - \ln(Z)\right),\end{aligned} \qquad [\mathbf{31}]$$

where Z is the normalization factor for the steady-state distribution. We find an expression for the derivatives $\frac{dP(i)}{dt}$ from the evolutionary dynamic. The evolutionary dynamic in continuous time can be written in terms of the master equation (3)

$$\frac{dP(i)}{dt} = \sum_{k\neq i}\left(W_{i,j}P(j) - W_{j,i}P(i)\right). \qquad [\mathbf{32}]$$

Because we know this dynamic satisfies DB, we can rewrite the master equation as

$$\frac{dP(i)}{dt} = \sum_{k\neq i} W_{i,j}P^*(j)\left(\frac{P(j)}{P^*(j)} - \frac{P(i)}{P^*(i)}\right). \qquad [\mathbf{33}]$$

Substituting this expression into Eq. **31** and reorganizing the sum, we get

$$\frac{dG}{dt} = \frac{1}{\nu}\sum_i\sum_{j\neq i} W_{i,j}P^*(j)\left(\frac{P(i)}{P^*(i)} - \frac{P(j)}{P^*(j)}\right)\left(\ln(\frac{P(i)}{P^*(i)}) - \ln(\frac{P(j)}{P^*(j)})\right). \qquad [\mathbf{34}]$$

Because $\ln(x)$ is a monotone function of x, each term in the sum is non-negative and therefore

$$\frac{dG}{dt} \geq 0. \qquad [35]$$

Incorporation of Asymmetric Mutation

Here we show that our formalism and the analogy with statistical physics can be generalized to incorporate asymmetry in mutation. To ground our discussion in biological reality, we consider two well established mutational structures: transition bias (4) and GC or AT bias (5). Transition bias implies that transitions, namely mutations within pyrimidine bases $Y = \{T, C\}$ or purine bases $R = \{A, G\}$, occur more frequently than transversions, i.e., mutations between these sets. Transition bias can be symmetric. For example, the following mutational scheme

$$\begin{aligned} &\mu_{T\to A} = \mu_{A\to T} = \mu_{T\to G} = \mu_{G\to T} = \mu_{C\to A} = \mu_{A\to C} = \mu_{C\to G} = \mu_{G\to C} = \mu \\ &\mu_{T\to C} = \mu_{C\to T} = \mu_{A\to G} = \mu_{G\to A} = \kappa\mu \end{aligned} \qquad [36]$$

where $\kappa > 1$ is the transition bias, is symmetric. A mutational scheme that generates GC or AT compositional bias is, however, necessarily asymmetric. For example, we consider the following mutational scheme, which incorporates both transition and compositional bias,

$$\begin{aligned} &\mu_{T\to A} = \mu_{A\to T} = \mu_{C\to G} = \mu_{G\to C} = \mu \\ &\mu_{A\to C} = \mu_{T\to G} = \mu \\ &\mu_{C\to A} = \mu_{G\to T} = c\mu \\ &\mu_{A\to G} = \mu_{T\to C} = \kappa\mu \\ &\mu_{G\to A} = \mu_{C\to T} = c\kappa\mu \end{aligned} \qquad [37]$$

where c is responsible for the compositional bias, such that $c > 1$ corresponds to AT bias, and $c < 1$ corresponds to GC bias. This scheme is asymmetric, because whenever $x \in \{A, T\}$ and $y \in \{G, C\}$, then

$$\frac{\mu_{x\to y}}{\mu_{y\to x}} = c. \qquad [\mathbf{38}]$$

The mutational scheme (Eq. **37**) is consistent with empirical data (for example, see ref. 6). Here we show that such a mutational scheme can be incorporated into the framework we have presented, leaving our derivations intact.

To generalize the formalism, we return to the focal point of our derivation (Eq. **5**). Our derivation requires that when we consider any two genotypes i and j, which are separated by a single mutation, the ratio of the transition rates between them can be expressed as

$$\frac{W_{j,i}}{W_{i,j}} = \frac{F'(j)}{F'(i)}, \qquad [\mathbf{39}]$$

where F' is a function of the state. In the mutational scheme (Eq. **37**), we see that if the transition between genotype i and j involves a point mutation within the base sets $\{A,T\}$ or $\{G,C\}$, then

$$\frac{W_{j,i}}{W_{i,j}} = \frac{F(j)}{F(i)}, \qquad [\mathbf{40}]$$

where F is the same function as in the symmetric case. Whereas if the transition between genotype i and j involves a point mutation between the base sets $\{A,T\}$ and $\{G,C\}$, for example, if the mutation from i to j involves a mutation from G to A, then

$$\frac{W_{j,i}}{W_{i,j}} = \frac{\mu_{G\to A}}{\mu_{A\to G}}\frac{F(j)}{F(i)} = c\frac{F(j)}{F(i)}. \qquad [\mathbf{41}]$$

If we can find a function F' such that relation **39** holds for any point mutation, then generalizing our formalism is straightforward.

A suitable function F' can be found from the analogy with statistical physics. In the mutation scheme, we have defined a mutation between sets $\{A,T\}$ and $\{G,C\}$ is analogous to a chemical reaction that involves a changes in energy, whereas a mutation within these sets is analogous to a reaction that does not involve such a change. Accordingly, we can express the asymmetry c in terms of the difference in "chemical potential" involved in a mutation from the set $\{A,T\}$ to the set $\{G,C\}$, namely,

$$c = e^{\nu(\mu_{AT}-\mu_{GC})}, \qquad [\mathbf{42}]$$

where ν is analogous to β, and μ_{AT} and μ_{GC} are analogous to the chemical potentials (7). This relation defines μ_{AT} and μ_{GC} up to an additive factor that does not affect the behavior of the system. Correspondingly, we can define F' as

$$F'(i) = \phi(i)F(i), \qquad [\mathbf{43}]$$

where $\phi(i)$ is defined as

$$\phi(i) = e^{\nu(\mu_{AT}n_{AT}(i)-\mu_{GC}n_{GC}(i))}, \qquad [\mathbf{44}]$$

where $n_{AT}(i)$ and $n_{GC}(i)$ are the numbers of AT and GC nucleotides in genotype i. It is easy to see that such a definition of F' satisfies condition **39** for any point mutation in the mutational scheme (Eq. **37**).

To complete this example, we must show how the results of the paper, namely the steady-state distribution, detailed balance, and free fitness, generalize under this mutation scheme. The new steady-state distribution is derived in the same way as in the symmetric case, namely,

$$P^*(i) = \frac{F'(i)}{\sum_k F'(k)} = \frac{e^{\nu(\ln(f_i)+\mu_{AT}n_{AT}(i)+\mu_{GC}n_{GC}(i))}}{Z}, \qquad [\mathbf{45}]$$

where

$$Z = \sum_k e^{\nu(\ln(f_k)+\mu_{AT}n_{AT}(k)+\mu_{GC}n_{GC}(k))}, \qquad [\mathbf{46}]$$

is the normalization factor. It is easy to verify that this steady-state solution satisfies DB. The free fitness is generalized in the same way chemical reactions are incorporated into the free energy, namely,

$$G = \langle \ln(f) \rangle + \mu_{AT} \langle n_{AT} \rangle + \mu_{GC} \langle n_{GC} \rangle + \frac{1}{\nu} H. \qquad [\mathbf{47}]$$

The proof that this is indeed a Lypunov function for the asymmetric mutation scheme is similar to that given above for the symmetric case.

We have shown that our framework can be generalized to incorporate the asymmetric mutation scheme (Eq. **37**). Such a generalization can incorporate up to three independent asymmetry parameters (in the above example, we had only one, which was c). This becomes apparent when we consider the thermodynamic interpretation of our generalization. We introduced asymmetry by partitioning the nucleotides into sets, where each set is associated with a given chemical potential. The asymmetry in mutation resulted from mutations that involved a change in chemical potential. We have three independent parameters at most, because we can partition four nucleotides into four sets at most, and only differences in chemical potential influence the asymmetries. In the most general case, a mutational scheme defined on four bases can have $\binom{4}{2} = 6$ independent asymmetries. It follows that only a subset of mutational schemes fall nicely into the generalized framework presented here. Nevertheless, this generalization already allows the incorporation of mutation schemes that are consistent with most of our knowledge about mutation in biology. Moreover, if mutation does not fall nicely into our framework, this implies only that the results presented here cannot be derived in a

transparent manner. However, based on studies in neural networks (8), our anticipation is that an asymmetric scheme can be approximated by a solvable one, such that the discrepancy will not affect most of the dynamic behaviours deduced based on the solvable scheme.

1. Crow, J. F. & Kimura, M. (1970) *An Introduction to Population Genetics Theory* (Harper & Row, New York).

2. Ewens, W. H. (1979) *Mathematical Population Genetics*, Levin, S. A., ed. (Springer, New York), Vol. 9.

3. van Kampen, N. G. (1981) *Stochastic Processes in Physics and Chemistry* (North–Holland, Amsterdam).

4. Freese, W. (1961) *Proc. Natl. Acad. Sci. USA* **47**, 540–545.

5. Bernardi, G. (1986) *J. Mol. Evol*. **24,** 1–11.

6. Petrov, D. A. & Hartl, D. L. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 1475–1479.

7. Reichl, L. E. (1980) *A Modern Course in Statistical Physics* (Univ. of Texas Press, Austin, TX).

8. Krogh, A., Hertz, J. & Palmer, R. (1991) *Introduction to the Theory of Neural Computation* (Addison–Wesley, New York).

Table 2. A comparison between the probabilities of fixation measured by simulations of the Wright-Fisher process (WF) and approximated by the canonical formula (CK) and the new formula (New)

N	Ns	S	WF *P* = 1/N	CK *P* = 1/N	New *P* = 1/N	WF *P* = 0.1	CK *P* = 0.1	New *P* = 0.1	WF *P* = 0.5	CK *P* = 0.5	New *P* = 0.5
100	-0.1	-0.001	0.00902 (3e-5)	0.00904	0.00904	0.09125 (9e-5)	0.09124	0.09124	0.4752 (0.0002)	0.4750	0.4750
100	-1	-0.01	0.00314 (2e-5) *	0.00316	0.00314	0.03443 (6e-5) *	0.03465	0.03444	0.2682(0.0001) *	0.2689	0.2680
100	-2	-0.02	0.000732 (9e-6) *	0.000761	0.000738	0.00895 (3e-5) *	0.00918	0.00891	0.1170 (0.0001)*	0.1192	0.1171
100	-5	-0.05	4e-006 (6e-7) *	4.8e-06	3.8e-06	6.0e-005 (2e-6)*	7.8E-05	6.3E-05	0.00579 (2e-5) *	0.00669	0.00589
100	-10	-0.1	0	4.6e-10	1.7e-10	0	1.3-08	5.1E-09	2.7e-005 (2e-6) *	4.59e-5	2.7e-5
100	0.1	0.001	0.01100 (35e-5)	0.01102	0.01102	0.1092 (0.0001)	0.1092	0.1092	0.5248(0.0002)	0.5250	0.5250
100	1	0.01	0.02279 (5e-5) *	0.02290	0.02282	0.2092 (0.0001)*	0.2096	0.2090	0.7299(0.0001) *	0.7311	0.7301
100	2	0.02	0.03957 (6e-5) *	0.03994	0.03959	0.3338 (0.0001)*	0.3358	0.3334	0.8788 (0.0001)*	0.8808	0.8787
100	5	0.05	0.09379 (9e-5) *	0.09517	0.09298	0.6256(0.0002) *	0.6321	0.6231	0.99256 (3e-5) *	0.99331	0.99245
100	10	0.1	0.1760 (0.0001) *	0.1813	0.1736	0.8552 (0.0001)*	0.8647	0.8514	0.999935 (3e-6)*	0.999955	0.999927
100	100	1	0.7954 (0.0001) *	0.8647	0.75	1	1	1	1	1	1
1000	-0.1	-0.0001	0.000905 (9.5e-6)	0.000903	0.000903	0.09110 (9e-5)	0.09124	0.09124	0.4748 (0.0002)	0.4750	0.4750
1000	-1	-0.001	0.000325 (6e-6)	0.000313	0.000313	0.03458 (6e-5) *	0.03465	0.03463	0.2691 (0.0001)!!	0.2689	0.2688
1000	-2	-0.002	7.6e-005 (3e-6)	7.5e-5	7.5e-5	0.00915 (3e-5) *	0.00918	0.00915	0.1191 (0.0001)	0.1192	0.1190
1000	-5	-0.005		4.6e-7	4.5e-7	7.7e-5 (3e-6)	7.80E-05	7.6E-05	0.00663 (3e-5) *	0.00669	0.00661
1000	-10	-0.01		4.2e-11	3.8e-11		1.38e-8	1.2e-8	4.0e-005 (2e-6) *	4.5e-5	4.3e-5
1000	0.1	0.0001	0.00109 (1e-5)	0.00110	0.00110	0.1091 (0.0001)	0.1092	0.1092	0.5252(0.0002)	0.5250	0.5250
1000	1	0.001	0.00228 (2e-5)	0.00231	0.00231	0.2094 (0.0001)	0.2096	0.2096	0.7308 (0.0001)*	0.7311	0.7310
1000	2	0.002	0.00403 (2e-5) *	0.00407	0.00406	0.3354 (0.0001)*	0.3358	0.3355	0.8805 (0.0001)*	0.8808	0.8806
1000	5	0.005	0.00989 (3e-5) *	0.00995	0.00993	0.6316 (0.0002)*	0.6321	0.6312	0.99319 (3e-5) *	0.9933	0.9932
1000	10	0.01	0.01967 (4e-5) *	0.01980	0.01970	0.8637 (0.0001)*	0.8647	0.8633	0.999953 (2e-6)*	0.999955	0.999952
1000	100	0.1	0.1761 (0.0001) *	0.1813	0.1736		1	1		1	1

Each fixation probability was measured using 107 simulated paths, and the standard deviations appear in parentheses. * indicates that the new formula is closer to the measured value than the canonical one, whereas !! indicates the opposite. In the only case where the canonical formula provides a better approximation to the measured value, the rounding misleadingly amplifies the difference: the measured value was 0.26908 (0.00014), and the canonical formula yielded 0.26894, whereas the new formula yielded 0.26884. Thus, this is probably the result of a sampling error. The general trend is not in any way amplified by rounding.