Directionality theory: an empirical study of an entropic principle in life-history evolution

Martin Ziehe[1] and Lloyd Demetrius[2,3,*]

[1]*Institut für Forstgenetik und Forstpflanzenzüchtung, Universität Göttingen, Büsgenweg 2,37089 Göttingen, Germany*
[2]*Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA*
[3]*Max Planck Institute for Molecular Genetics, Ihnestraße 71, 14195 Berlin, Germany*

Understanding the relationship between ecological constraints and life-history properties constitutes a central problem in evolutionary ecology. Directionality theory, a model of the evolutionary process based on demographic entropy, a measure of the uncertainty in the age of the mother of a randomly chosen newborn, provides an analytical framework for addressing this problem. The theory predicts that in populations that spend the greater part of their evolutionary history in the stationary growth phase (equilibrium species), entropy will increase. Equilibrium species will be characterized by high iteroparity and strong demographic stability. In populations that spend the greater part of their evolutionary history in the exponential growth phase (opportunistic species), entropy will decrease when population size is large, and will undergo random variation when population size is small. Opportunistic species will be characterized by weak iteroparity and weak demographic stability when population size is large, and random variations in these attributes when population size is small. This paper assesses the validity of these predictions by employing a demographic dataset of 66 species of perennial plants. This empirical analysis is consistent with directionality theory and provides support for its significance as an explanatory and predictive model of life-history evolution.

Keywords: Malthusian parameter; demographic entropy; Darwinian fitness; demographic stability; plant populations

1. INTRODUCTION

One of the main challenges in evolutionary ecology is to explain the match between ecological conditions, as agents of selection, and the demographic characteristics of the organisms that are selected under such conditions. In the 1970s, the model of $r - K$ selection emerged as an influential response to this challenge (Pianka 1970, after MacArthur & Wilson 1967). Although this model provided some qualitative insight into the relationship between ecological constraints and life-history characteristics, its weakness as a predictive model is now recognized (Stearns 1992; Reznick *et al.* 2002). These shortcomings derive largely from the fact that the $r - K$ selection model (in its analytic expression as distinct from its more qualitative claims) is essentially concerned with populations in which fertility and mortality variables are independent of age. Accordingly, the model is unable to explain in quantitative terms the correlation between ecological conditions, such as density-dependent constraints, and age-dependent life-history characteristics, such as age of sexual maturity, reproductive span and longevity.

Given the limitations of the $r - K$ model, we may ask whether demographic structure and Mendelian genetics can be integrated in an evolutionary context, which would explain the empirical relations between ecological conditions and life history. This problem has generated a large literature in ecological and demographic genetics, as evidenced by the contributions of Crow (1979), Lande (1982) and Charlesworth (1994), among others.

The central problem that needs to be addressed in these studies is determining the class of population parameters which predicts the outcome of competition between a variant type and the incumbent in populations structured by age. In other words, what function of the age-specific fecundity and mortality variables characterizes Darwinian fitness?

In models that lack a demographic structure, individuals are assumed to be indistinguishable in terms of their age. Fitness is determined by counting the number of individuals in the population and assessing the rate of increase—the Malthusian parameter, denoted r. In this class of models, the invasion of a mutant allele is predicted by the relative values of r. This characterization of Darwinian fitness forms the core of a large body of work in classical population genetics (e.g. Crow & Kimura 1970), and constitutes the basis of the $r - K$ selection paradigm.

In models defined by demographic structure, individuals are assumed to differ in terms of their age, size or some other morphometric or physiological characteristic. Fitness in this context can also be described by counting the number of individuals in the population and assessing the rate of increase when the stable age distribution is attained (Fisher 1958). This growth rate parameter is now a function of the age-specific fecundity and mortality variables, and is described by the real root of the Lotka equation

$$1 = \int_0^\infty \exp(-rx)V(x)\mathrm{d}x. \tag{1.1}$$

Here, $V(x)$ is the net fecundity function at age x, and is defined by the product $V(x) = l(x)m(x)$, where $l(x)$ is

the probability that an individual born at age zero survives to age x, and $m(x)$ the mean number of offspring produced by an individual at age x.

Studies based on the theory of branching processes (Charlesworth & Williamson 1975; Pollak 1976) gave rise to the claim that the outcome of competition between a mutant type and the incumbent is predicted by relative values of the Malthusian parameter, as defined by the Lotka equation. In view of this claim and the pioneering analytical studies in Fisher (1958), the Malthusian parameter emerged as a cornerstone in efforts to integrate demographic structure with Mendelian genetics in evolutionary dynamics.

Although the Malthusian models do provide some qualitative insight into life-history evolution, they have very weak predictive or explanatory power, both in an ecological setting, where the problem concerns the dynamics of invading species, and in an evolutionary context, where the problem concerns the dynamics of invading genotypes. The limitations of the model are evident from empirical observations on invasion studies (Lawton & Brown 1986) and from analytic studies of invasion in genetic models (Demetrius & Gundlach 1999; Demetrius 2001).

Using data from a wide variety of vertebrate and invertebrate populations, Lawton & Brown (1986) have observed that the probability of the establishment of an invader is positively correlated with body size, a variable which (as shown by allometric relations) is negatively correlated with the population growth rate (Calder 1996). These empirical studies thus question the significance of the Malthusian parameter as the main determinant of invasion success.

In studies of invasion in genetic models using diffusion processes, Demetrius & Gundlach (1999, 2000) showed that the probability of establishment of an invader, in contrast to the claims cited earlier, is not predicted by the Malthusian parameter but is contingent on other demographic variables. The studies described in Charlesworth & Williamson (1975) and Pollak (1976) pertain to a population size that is effectively infinite, and they consider the kinetics of invasion as a deterministic process. However, as shown in Demetrius & Gundlach (2000), when size is finite, the effect of demographic stochasticity on the population dynamics implies that random deviations from the stable age distribution will always occur. Accordingly, the kinetics of invasion now becomes a stochastic process, which involves not only the Malthusian parameter, but also other macroscopic variables such as the demographic variance; that is, the variance in the age of reproducing individuals in the population. The selective advantage, denoted by s, of a mutant type will now be given by

$$s = \Delta r - \frac{1}{N}\Delta\sigma^2. \tag{1.2}$$

Here, N denotes the population size, r the Malthusian parameter, as defined by equation (1.1), and σ^2 the demographic variance, a function of the age-specific fecundity and mortality variables. The quantities Δr and $\Delta\sigma^2$ refer to the differences in the life cycle parameters r and σ^2, respectively, between the mutant and the wild-type.

Equation (1.2) implies that the selective advantage s, as given by $s = \Delta r$, is only valid when $N \to \infty$. Consequently,

in structured populations, the Malthusian parameter characterizes Darwinian fitness only in the limiting case when the population size is effectively infinite.

The selective advantage given by equation (1.2) implicates several demographic variables. The problem of characterizing Darwinian fitness in structured populations in terms of a single demographic variable was finally resolved by appealing to a new class of mathematical ideas that originated from ergodic theory (Demetrius & Gundlach 1999, 2000). We recognized that in populations with demographic structure, the contribution of a type to successive generations involves assigning weights to each age class according to the frequency of their genealogies, a demographic concept that refers to the recording of successive ancestral states of a particular individual, which, at time 0, is in the first age class. This assignment of weights led to a new measure of population size, which we called the 'demographic effective size' (Demetrius 1983). We showed that the rate of increase of this effective size is a function of the age-specific fecundity and mortality variables, a parameter we call 'demographic entropy' because of its formal relation to the Boltzmann–Gibbs entropy in statistical thermodynamics. Furthermore, we showed that in studies of invasion dynamics, demographic entropy predicts the outcome of competition between a rare mutant and an incumbent population. We also observed that as the population size becomes effectively infinite, the demographic effective size (which weights individuals according to the frequency of their genealogies) and the census size (which assigns equal weight to each individual in the population) coincide. Hence, in the infinite size limit, the invasion criterion based on entropy reduces to the criterion based on the Malthusian parameter (Demetrius 2001). Consequently, in structured populations—of finite and effectively infinite size—demographic entropy characterizes Darwinian fitness.

Analytically, demographic entropy, denoted H, can be expressed in the form

$$H = -\frac{\int_0^\infty p(x)\log p(x)\,\mathrm{d}x}{\int_0^\infty xp(x)\,\mathrm{d}x} \equiv \frac{S}{T}. \tag{1.3}$$

The function $p(x)\mathrm{d}x$ represents the probability that the mother of a randomly chosen newborn belongs to the age interval $(x, x+\mathrm{d}x)$. The function $p(x)$ is given by

$$p(x) = \exp(-rx)V(x). \tag{1.4}$$

The quantity S is a pure number. It represents the uncertainty in the age of the mother of a randomly chosen newborn. S is a measure of the variability in the age at which individuals reproduce and die. The parameter T is the mean age of mothers at the birth of their offspring. The quantity H has dimensions of inverse of time; it is the rate of increase of the demographic effective size.

Since entropy predicts the outcome of competition between a rare mutant and an incumbent population, the selective advantage s of a mutant allele can now be expressed by

$$s = \Delta H. \tag{1.5}$$

As shown in Demetrius (2001), invasability is predicted by entropy. However, whether an increase or decrease in entropy confers a selective advantage will be contingent on two classes of ecological factors:

(i) *constraints on the population growth rate*: bounded growth ($0 \leq r < H$), unbounded growth ($r > H$);
(ii) *constraints on the population size*: large ($N > \Delta\sigma^2/\Delta r$), small ($N < \Delta\sigma^2/\Delta r$). These conditions derive from equation (1.2).

Under bounded growth conditions, increased entropy will almost always confer a selective advantage. When unbounded growth conditions prevail, decreased entropy will almost always confer a selective advantage when population size is large. However, its selective effect will be highly stochastic when population size is small.

The invasion criteria expressed in terms of H are given in table 1 (Demetrius 2001).

Analytic studies of properties of the parameters H and S show that evolutionary changes in the two entropic parameters are positively correlated (Demetrius 2000). This implies that the analytic results on evolutionary changes derived for the rate parameter H will also pertain to the information measure S.

The significance of entropy in a demographic or ecological context derives from the fact that it completely characterizes the robustness of a population; that is, the capacity of macroscopic parameters of the population to maintain their steady state values in the face of random perturbations in the age-specific fecundity and mortality variables. This stability property, which emerged from a series of analytical studies (Demetrius 1977; Tuljapurkar 1982; Demetrius *et al.* 2004), can be qualitatively expressed as follows.

1 Entropy is positively correlated with the rate of decay of fluctuations in population numbers. Accordingly, entropy provides a precise measure of demographic stability; that is, the ability of a population to maintain its predicted trajectory in spite of perturbations induced by intrinsic or extrinsic factors.

The pertinence of entropy for evolutionary studies revolves around the fact that it predicts the outcome of competition between an invading mutant and a resident population. In view of this condition, entropy can be used to parametrize the direction of evolutionary change under mutation and natural selection. The mathematical analysis of this evolutionary process, which we call directionality theory, is concerned with changes in the genotypic and phenotypic composition of a population as one population type replaces another through the mutation-selection regime.

Directionality theory recognizes that the invasion criteria of a mutant are contingent on whether the population is stationary or fluctuating around some constant size, or increasing exponentially. Accordingly, the theory appeals to these ecological constraints to distinguish between two comparable classes of demographic constraints which may characterize a population throughout its evolutionary history, namely: (i) bounded growth—this condition pertains to populations (equilibrium species) that spend the greater part of their evolutionary history in the stationary growth phase; (ii) unbounded growth—this situation refers to populations (opportunistic species) that spend the greater part of their evolutionary history in the exponential growth phase.

Table 1. Invasion criteria for a rare allele. a.s., almost surely.

ecological constraints	demographic condition	selective outcome
bounded growth	$\Delta H > 0$	mutant invades (a.s.)
	$\Delta H < 0$	mutant becomes extinct (a.s.)
unbounded growth (large size)	$\Delta H < 0$	mutant invades (a.s.)
	$\Delta H > 0$	mutant becomes extinct (a.s.)
unbounded growth (small size)	$\Delta H < 0$	mutant invades with a probability increasing in n
	$\Delta H > 0$	mutant becomes extinct with a probability increasing in n

Directionality theory exploits the invasion criteria in table 1, which describe *local* changes in entropy to predict *global* changes in entropy as one population replaces another under the mutation-selection process. The analytical basis of the relations between ecological norms and long-term trends in entropy are reviewed in Demetrius *et al.* (2004). These relations can be qualitatively annotated as follows.

2(*a*) Under bounded growth conditions, evolution results in a unidirectional increase in entropy.
2(*b*) Under unbounded growth conditions and large population size, evolution results in a unidirectional decrease in entropy.
2(*c*) Under unbounded growth conditions and small population size, evolution results in a random, non-directional change in entropy.

This paper exploits plant populations to evaluate the empirical support for the entropy-stability tenet as described by (1), and the directionality principles for entropy as expressed by (2).

Empirical support for the relation between demographic stability and entropy, and the directionality principles for entropy, have been addressed in Kim & Schoen (1993), and in Demetrius & Ziehe (1984), respectively, by appealing to life-history data of human populations. Kim & Schoen (1993) exploited certain novel characterizations of demographic entropy from Tuljapurkar (1982) to show empirically that entropy predicts demographic stability. Demetrius & Ziehe (1984) analysed demographic data drawn from Sweden (1778 to 1965) and France (1851 to 1965) to study the correlation between ecological constraints and evolutionary changes in entropy.

Demographic and evolutionary studies based on human populations have the advantage that the database is highly reliable, and the ecological constraints can be precisely delineated. However, these studies have the disadvantages that variations in the ecological factors are restrained, and that changes in the demographic variables are due primarily to cultural forces. In this paper, we obviate these disadvantages by focusing our study on plant populations. We draw on the plant demography database compiled by Silvertown and Franco since the early 1990s (e.g. Franco & Silvertown 1996). The computation of

entropy and other life-history variables given in the Electronic Appendix was done by Miguel Franco. Our empirical analysis of the entropy-stability tenet and the directionality principles for entropy will appeal to this computational study.

The paper is organized as follows. In §2 we provide a brief account of the origin of the entropy concept and review the analytical basis for the characterization of entropy as a measure of demographic stability. The main ideas that underlie the directionality principles are described in §3.

The empirical analysis is developed in §§5 and 6. Section 5 uses the life-history database to analyse the relation between entropy and demographic stability, and §6 is concerned with the empirical support for the directionality principles.

We refer readers to the companion article by Kowald & Demetrius (2005), which gives a computational study of the entropic principles assessed in this paper.

2. ENTROPY: ORIGIN AND PROPERTIES

The concept demographic entropy, as defined by the expression S in equation (1.3), describes the variability in the age at which individuals in a population reproduce and die. This notion of demographic variability is a fundamental property of all populations of replicating organisms. The property has its origin in the processes that underlie the ontogeny of the individual. In cellular systems it results from the random inequalities between cells, such as the unequal distribution of metabolic components, which occur at cell division. In multicellular and higher organisms, demographic heterogeneity derives from the small variations in the sequence of developmental events that transform the zygote into an adult. Accordingly, any genetically homogeneous population of organisms will be characterized by heterogeneity in its life cycle, and by variability in the rates at which individuals in the population reproduce and die.

Demographic heterogeneity thus constitutes an intrinsic property that must be considered in any theory designed to explain the persistence of a population and its adaptation to the environmental conditions. The problem that now arises is to determine the class of mathematical objects which will characterize demographic heterogeneity, and also provide a basis for predicting the adaptive state or robustness of a population.

The resolution of this problem is confounded by the fact that individual birth and death rates—the variables that are integrated to predict the adaptive state of a population—can be parametrized in terms of different individual states, including age, size or metabolic energy. Different parametrizations will lead to different mathematical models of the demographic process and may thus yield different values for the life-cycle heterogeneity. A measure of demographic heterogeneity, such as the variance in the net reproductive function, may yield one value in a model parametrized by age and another value in a model parametrized by size, even though the age and size classes were defined in such a way that the two models were dynamically equivalent.

In order to have predictive power, an index of life-cycle heterogeneity should be relatively independent of the variables—age, size or energy—in which the individual birth and death rates are expressed. In a more formal language, the heterogeneity index should be invariant with respect to the mode of parametrization invoked. Different models of the demographic process, if dynamically equivalent, should be described by similar values for their life-cycle heterogeneity.

In Demetrius (1974), we appealed to developments in the ergodic theory of dynamical systems (Billingsley 1965; Sinai 1989) to show that there is a unique measure of heterogeneity that is an invariant of the class of models describing the demographic process. This demographic invariant is the dynamic entropy (the Kolmogorov–Sinai entropy) of the statistical process generated by the demographic model.

In systems where the parametrization of the demographic process is in terms of age, rather than size or metabolic energy, the demographic model has a relatively simple representation—the Leslie model (in discrete time) and the Lotka model (in continuous time). For these systems, the Kolmogorov–Sinai entropy can be computed by representing the demographic model as a Markov process. In the case of a continuous age-structured model, the demographic entropy is given by H, as defined in equation (1.3). The numerator in equation (1.3) is the information measure S, given by

$$S = -\int_0^\infty p(x)\log p(x)\mathrm{d}x, \qquad (2.1)$$

where the function $p(x)$ is described by equation (1.4).

The demographic parameter S is a measure of iteroparity and is determined by the net reproductive function $V(x)$. The values that S assumes depend on the shape or degree of concavity of the net reproductive function.

Populations characterized by large and small values of S can be distinguished in terms of the following life-history parameters: age of sexual maturity, reproductive span and litter size. Studies described in Demetrius (2004) based on the sensitivity of entropy to changes in the age-specific fecundity and mortality variables have delineated the following relations: (a) large S—late age of sexual maturity, broad reproductive span and small litter size; (b) small S—early age of sexual maturity, narrow reproductive span and large litter size.

Figure 1a,b illustrates these two entropic states with net reproductive functions that are representative for a tree (a high entropy species) and a herb from a disturbed habitat (a low entropy species).

In figure 1a, the parametrization is in terms of 10-year age classes. The entropy S is given by $S=2.24$; the generation time is $T=84$ years. In figure 1b, we use 1-year age classes. Here, $S=0.95$ and $T=2.5$ years.

(a) *Entropy and demographic stability*

Changes in birth and death rates owing to intrinsic factors (demographic stochasticity) or extrinsic forces (environmental stochasticity) are a characteristic property of all natural populations. In view of these changes, the observed values of macroscopic population variables will deviate from their predicted values. This property entails that parameters such as the population growth rate will undergo fluctuations from their steady state values. Populations will differ in terms of the magnitude of their deviations from the steady state-condition, and
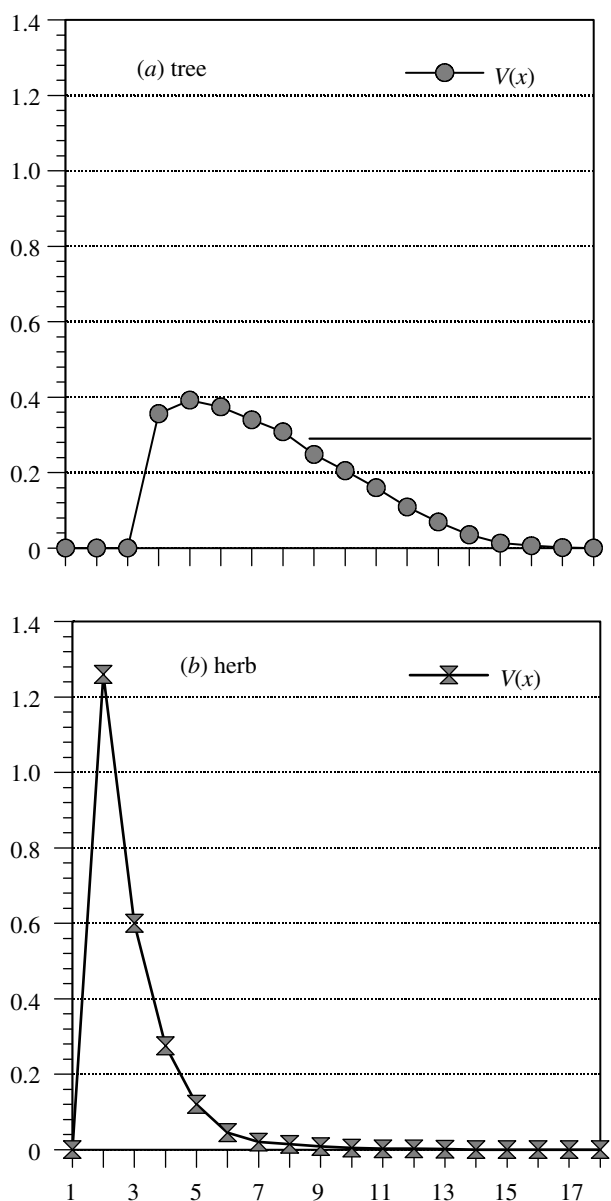
Figure 1. Schematic net reproductive functions $V(x)$ representative for (a) a tree and (b) a herb.

also the rate at which these fluctuations from the steady-state condition are attenuated. Strongly stable populations are defined by small deviations, and weakly stable populations by large deviations from the steady-state condition. These observations led to the hypothesis that the degree of demographic stability can be characterized by the fluctuation decay rate of macroscopic observables towards their steady state values. These qualitative ideas were given a mathematical representation by appealing to the ergodic theory of population processes (Demetrius *et al.* 2004).

Analytically, the fluctuation decay rate can be described as follows. Let $P_n(\varepsilon)$ denote the probability that the population size at instant n differs from the predicted value by more than ε. As n increases, it can be shown that $P_n(\varepsilon)$ will tend to zero. The fluctuation decay rate, Q, is the asymptotic rate at which the number $P_n(\varepsilon)$ tends to zero as n tends to infinity. The quantity $P_n(\varepsilon)$ can be interpreted as the probability that the sample mean defined at instant n differs from the asymptotic mean by more than ε.

Analytically, we write

$$Q = \lim_{n \to \infty} \left[-\frac{1}{n} \log P_n(\varepsilon) \right]. \tag{2.2}$$

An analogous characterization of the fluctuation decay rate was given in Demetrius (1977) in an effort to address certain problems raised by Coale (1972) and Keyfitz (1972) in their studies of fluctuation intensity in human populations, Demetrius (1977) invoked this analogous measure of fluctuation intensity to provide a partial analytical support for the hypothesis that demographic entropy is positively correlated with demographic stability.

The complete analytical support for the entropy-stability hypothesis was recently given in Demetrius *et al.* (2004) in terms of the fluctuation decay rate concept defined by equation (2.2). We appealed to the theory of large deviations to establish a fluctuation theorem, which asserts that the entropy H and the fluctuation decay rate Q are positively correlated.

The analytical formulation is given by

$$\Delta H \Delta Q > 0. \tag{2.3}$$

Here, ΔH and ΔQ represent changes in the macroscopic variables H and Q, respectively, induced by a change in the age-specific fecundity and mortality variables.

Equation (2.3) essentially asserts that the changes in H and Q, induced by small changes in the age-specific fecundity and mortality variables, are positively correlated. In the case of an evolutionarily stable population, that is, a population whose life-table is defined by extremal values of entropy (maxima for equilibrium species, minima for opportunistic species), equation (2.3) implies that H and Q will also be positively correlated.

We should emphasize that, in populations that are far removed from the evolutionarily stable condition, the entropy H and the fluctuation decay rate Q need not be positively correlated. However, the entropy function, defined by the parameter S, and the fluctuation decay rate Q, a quantity which will be positively correlated with the physiological variable metabolic rate, both satisfy certain allometric relations in terms of body size (Demetrius 2000). These relations entail that S and Q will always be positively correlated. We write

$$\Delta S \Delta Q > 0. \tag{2.4}$$

The equation (2.4) asserts that the greater the degree of iteroparity, the more insensitive are the population observables to perturbations in the genetic and environmental parameters. In other words, a high degree of iteroparity implies an enhanced robustness or resilience to perturbations in the age-specific fecundity and mortality variables.

3. DIRECTIONALITY THEORY
Biological evolution (the change in diversity and adaptation of populations over time) involves two complementary processes—mutation and selection. Mutation generates genetic variability. Selection orders this variability through competition between the ancestral and mutant types. Directionality theory invokes the fundamental attributes of demographic entropy as a measure of Darwinian fitness, and studies the changes in genotypic and phenotypic composition that are generated by the mutation-selection regime.

The models recognize that populations can be classified in terms of the amount and availability of the resources. The theory distinguishes between resources that are constant but limited, and resources that are ample but intermittent. Populations subject to constant but limited resource conditions (equilibrium species) will typically be characterized by a growth rate that is either stationary ($r = 0$) or fluctuates around this stationary value. Populations that are subject to ample but intermittent resources (opportunistic species) will typically be described by periodic bursts of exponential increase. These two extremes of environmental conditions—representative of the conditions of bounded and unbounded growth described earlier—can be distinguished by means of a demographic parameter called the reproductive potential, denoted by Φ, and given by

$$\Phi = \frac{\int_0^\infty p(x)\log V(x)\mathrm{d}x}{\int_0^\infty xp(x)\mathrm{d}x} \equiv \frac{E}{T}. \tag{3.1}$$

The numerator E corresponds to the offspring production rate, $\log V(x)$, averaged over the different age classes.

The quantities r, H and Φ are related by the identity

$$\Phi = r - H. \tag{3.2}$$

We derive from equation (3.2) that the following implications hold:

$$\Phi \leq 0 \Rightarrow r \leq H \quad \text{and} \quad \Phi > 0 \Rightarrow r > H. \tag{3.3}$$

The condition $\Phi < 0$ (bounded growth) describes populations whose growth rate is bounded by entropy. It analytically characterizes the notion of a population subject to limited but fairly constant resource constraints. The condition $\Phi > 0$ (unbounded growth) describes populations with a growth rate that exceeds entropy. Since resources are finite, this state of rapid exponential growth cannot continue indefinitely, as resources will ultimately become depleted. The unbounded growth condition thus describes the exponential growth phase of a population subject to intermittent episodes of ample resource constraints.

It is of some interest to note that, by observing that $E/T \equiv \Phi$ and $S/T \equiv H$, equation (3.2) can be written in the form

$$E = rT - S \tag{3.4}$$

and we have

$$E < 0 \Rightarrow rT < S; \quad E > 0 \Rightarrow rT > S. \tag{3.5}$$

The relations (3.2) and (3.4), and the corresponding implications (3.3) and (3.5), underscore the distinction between the entropy parameter H, which has the dimension of inverse time and describes a rate r, and the parameter S, a pure number which describes simply the variability in age-specific reproduction and mortality.

(a) *Directionality principles*
The directionality principles described by 2(a), 2(b) and 2(c) are formulated in terms of relations between ecological constraints, as defined by the bounded ($\Phi < 0$) and unbounded ($\Phi > 0$) growth distinctions, and the changes in the entropy function H as one population replaces another under the mutation-selection process.

Since $E = \Phi T$, we can express the entropic principles in terms of correlations between the ecological norms as

defined by E, and the directional trends in entropy, S. We have

3(a) $E < 0$: a unidirectional increase in S.

3(b) $E > 0$, large population size: a unidirectional decrease in S.

3(c) $E > 0$, small population size: random, non-directional change in S.

The empirical study we develop will be based on the above correspondence between ecological norms and the long-term changes in entropy.

4. EMPIRICAL CONSIDERATIONS: METHODS
The data we used to test the model are from the plant demography database compiled by Franco & Silvertown (1996). The values for the demographic parameters are the result of a computation analysis by Miguel Franco using the demographic data base. Matrix projection models for stage data of the species listed in the Electronic Appendix were employed to calculate life tables using the program STAGE-COACH (Cochran & Ellner 1992). These life tables were then used to calculate the parameters: growth rate, entropy, reproductive potential, lifespan and generation time, described in the equations presented here (see tables in the Electronic Appendix). Because of rounding errors, the intrinsic rate of increase produced by matrix projection differed slightly from that obtained employing Lotka's equation (equation (1.1)) on the projected life tables. Therefore, the latter was used in all the analyses. In addition, because semelparous (monocarpic) plants always yield $S = H = 0$, only iteroparous perennials were considered.

With regard to the statistical analysis of the data, we employed cross-species comparisons and not phylogenetically independent contrasts (see Harvey & Pagel 1991) for two reasons. First, the work of Franco & Silvertown (1996) employing this database has shown that because the species come from a wide variety of families, independent contrasts and cross-species analyses produce similar results. Second, the use of cross-species comparisons allows the reader to visualize the patterns discussed in terms of the precise values of the parameters.

5. ENTROPY AND STABILITY: EMPIRICAL ANALYSIS
The entropy-stability principle expressed by equation (2.4) asserts that the entropy function S is positively correlated with demographic stability as defined by the fluctuation decay rate Q. The entropy function S is explicitly given in terms of the age-specific fecundity and mortality variables. Hence, S can be derived from the life-table data. Table 3 in the Electronic Appendix gives the values of S for the 66 species considered. The 66 species fall into four categories: trees (t), shrubs (s), forest herbs (f) and herbs from disturbed habitats (o). Although a large variation exists within the different groups, these groups can be ordered in terms of decreasing entropies S:

trees (t) > shrubs (s) > forest herbs (f)
> herbs from disturbed habitats (o).

The fluctuation decay rate Q is defined in terms of an asymptotic limit (see equation (2.2)). Hence, Q cannot be assessed directly from the life-table data. However,

the fluctuation decay rate is inversely related to the intensity of fluctuations. Hence, in principle, Q can be estimated by considering the observable population size, a demographic parameter whose fluctuation intensity can be estimated from time-series data. The problem of estimating the fluctuation decay rate thus reduces to the existence and reliability of these data. In plant populations, unlike human and laboratory populations, precise data on population size and its temporal variation are not generally accessible. To our knowledge, the only studies that exist pertain to qualitative observations of a comparative nature. These general observations in the case of the 66 species under study permit an ordering of the four groups of plants in terms of increasing intensity of fluctuations in population numbers, leading to exactly the opposite ranking as for the entropy S. We can infer from these considerations that the relation between fluctuation intensity and entropy, as revealed by the empirical plant data, is consistent with the entropy-stability hypothesis.

We can build on these observations to generate a quantitative test of the entropy-stability principle. Our analysis will be based on a widely held ecological tenet: fluctuation intensity in population size is inversely related to generation time. This tenet has sometimes been expressed in terms of correlations between fluctuation intensity in population numbers and the morphometric variable body size (Emlen 1984, ch. 2). Since body size is positively correlated with generation time (Bonner 1988, ch. 2), the two tenets are equivalent.

We can integrate this fluctuation intensity-generation time tenet with the entropy-stability principle, as expressed by equation (2.4), to predict that entropy and generation time will be positively correlated. We write

$$\Delta S \Delta T > 0. \tag{5.1}$$

Since entropy and generation time can be evaluated from the life-table data, we can now assess the entropy-stability tenet by using the values for S and T given in table 3 in the Electronic Appendix.

The relation between S and T is given in figure 2. We observe that S and T are positively correlated, which is consistent with the prediction of the entropy-stability principle.

In the graph, we have distinguished between the four groups of plants in order to examine the nature of the correlation within each group. We observe that among trees, the relation between S and T is approximately linear. This linearity has an evolutionary rationale. Among the four groups of plants, trees are the prototypic equilibrium species. Hence, evolution in this group will be described by a unidirectional increase in entropy, and concomitantly, a unidirectional increase in generation time. The life history of this group of plants will be described by states that are close to the conditions described by maximal entropy and maximal generation time. The linear correlation between S and T is a consequence of the fact that both parameters will be close to their extremal values owing to evolutionary considerations.

6. DIRECTIONAL CHANGES IN ENTROPY: AN EMPIRICAL STUDY

Ecological constraints have been expressed in terms of the parameter E defined by $E = \Phi T$, where Φ denotes
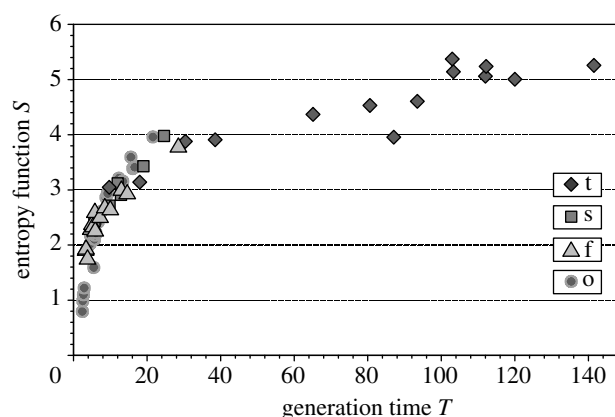
Figure 2. Relationship between entropy, S, and generation time, T (years). (t, trees; s, shrubs; f, understory forest herbs; o, herbs from open disturbed habitats.)

the reproductive potential and T the generation time. The relations between ecological norms as defined by E, and changes in the entropy function S are given by $3(a$–c).

The directional trends in entropy refer to changes within a given phyletic lineage. These changes can be invoked to infer correlations between the values of E and S for the species among a diverse phylogenetic tree. We have:

A(1) $E < 0$: the smaller the value of E, the larger the value of S, that is, large values of S will be correlated with stronger negative values of E.

A(2) $E > 0$:
  (i) for populations with large size, the greater the value of E, the smaller the value of S; that is, small values of S will correspond to large positive values of E.
  (ii) for populations of small size, the values of E and S are uncorrelated.

The empirical studies we now evaluate will be based on the above correlations.

(a) *Empirical analysis: qualitative considerations*
Austad (1997) presented a comprehensive study of the relation between ecological constraints and life-history patterns in mammalian species that provide qualitative support for A(1) and A(2). The ecological constraints are described in terms of the notion of environmental hazard. Low environmental hazard corresponds to low risks of death from extrinsic factors such as predation and extreme climatic conditions, and high environmental hazard refers to a high risk of death due primarily to external constraints. The life-history patterns are represented in terms of demographic parameters such as age of sexual maturity, litter size and reproductive span.

The cross-species comparisons of the mammalian lineage, described in Austad (1997), delineate the following correlations between environmental hazard and demographic parameters.

B(1) *Low environmental hazard*: species with late age of sexual maturity, small litter size and broad reproductive span.

B(2) *High environmental hazard*: species with early age of sexual maturity, large litter size and narrow reproductive span.

We will show that these empirical observations are consistent with the entropic principles as expressed by A(1) and A(2). In order to establish this consistency, we will reformulate the empirical correlations between environmental hazard and life-history variables in terms of correlations between the parameters: average net offspring production rate E, and the entropy function S.

(i) Environmental hazard and the normalized reproductive potential

We will now show that $E < 0$ corresponds to low environmental hazard, and $E > 0$ to high environmental hazard. The argument underlying this correspondence is as follows. We note that $rT \sim \log_e R_o$, where R_o is the net reproductive rate given by $R_0 = \int_0^\infty V(x) dx$. From equation (3.4), we conclude that $E < 0$ is described by the condition $0 \leq \log_e R_o < S$; that is, a net reproductive rate, $R_o = 1$, or a rate whose logarithm is bounded by entropy. This growth rate condition describes a population that is constant in size or fluctuates around some constant value. This constraint on size is consistent with a mortality rate driven primarily by intrinsic factors, with low risks of death from extreme environmental conditions and thus low environmental hazard.

From equation (3.4), we also infer that the relation $E > 0$ corresponds to the condition $\log_e R_o > S$. This describes a net reproductive rate, whose logarithm exceeds the entropy. In this case, population size increases exponentially. As this state cannot continue indefinitely, the condition $E > 0$ will represent the exponential growth phase in a population characterized by recurrent episodes of population increase and sharp decline. This pattern of population growth is consistent with a high risk of death from extrinsic factors and concomitantly high environmental hazard.

(ii) Life-history patterns and entropy

Age of sexual maturity, litter size and reproductive span constitute three canonical variables. These three life-history parameters characterize demographic entropy (Demetrius 2004). High entropy is correlated with late age of sexual maturity, small litter size and broad reproductive span, whereas small entropy is described by early age of sexual maturity, large litter size and narrow reproductive span.

(iii) Environmental patterns and the entropic principle

Since $E < 0$ and $E > 0$ correspond to low and high environmental hazards, respectively, and entropy S characterizes the range of variations in the life-history variables age of sexual maturity, litter size and reproductive lifespan, we can infer that the empirical correlations expressed by B(1) and B(2) are equivalent to the analytical tenets described by A(1) and A(2). This equivalence entails that the empirical observations regarding environmental hazard and life-history patterns are consistent with the predictions of directionality theory that relate ecological constraints with entropy.

(b) Empirical analysis: quantitative considerations

Any quantitative evaluation of the entropic principles requires studies that incorporate the demographic parameters age-specific fecundity and mortality. The demographic data may be (i) longitudinal, where life-history parameters represent the demographic state of a single population at different points in evolutionary time; or
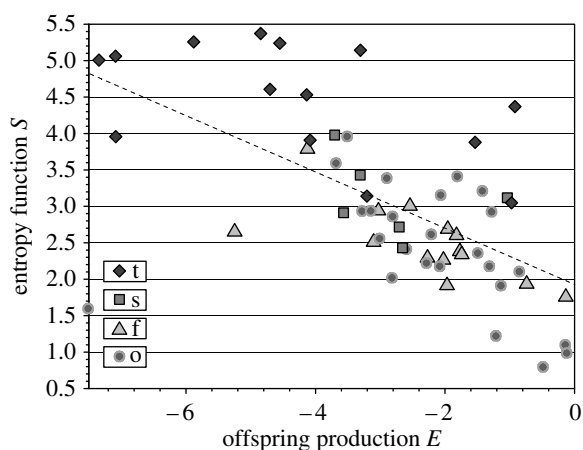

Figure 3. The relation between E and S for species that satisfy the condition $E < 0$ (equilibrium species). (Letter codes as in figure 2.)

(ii) cross-sectional, where the life-history parameters describe the demographic state of several different species at the same point in evolutionary time.

The demographic parameters described in the Electronic Appendix provide cross-sectional data for 66 species. These data can be used to derive quantitative expressions for E and S, and thereby to quantitatively assess the predictions of directionality theory.

The plant data considered are drawn from a highly heterogeneous group—a situation that may complicate the interpretation of cross-species comparisons. However, the 66 species fall into four categories that are well-defined in terms of their phylogenetic status.

The species can also be classified in terms of the equilibrium $(E < 0)$/opportunistic $(E > 0)$ distinction we have proposed. Equilibrium species consist of 59 members. We empirically assess the relations between E and S for the complete set of species, and also for the four subgroups defined in terms of their phylogenetic status. Opportunistic species comprise only six members. In view of the small size of this group, we will evaluate the empirical data without considering any phylogenetic distinction.

(i) Equilibrium species (E < 0)

The directionality principles predict that the parameters E and S will be negatively correlated. We have evaluated the correlation coefficient between the parameters E and S for the four categories we have described: trees (t), shrubs (s), herbs from open disturbed habitats (o) and understorey forest herbs (f). The coefficients are all negative: understorey forest herbs (f): -0.6933; trees (t): -0.5258; herbs from open habitats (o): -0.3461; shrubs (s): -0.3409; all species: -0.6179.

Figure 3 illustrates the relations between E and S for the complete set of equilibrium species. For the four groups, the corresponding figures, which have the same pattern as in figure 3, can be found in the Electronic Appendix.

(ii) Opportunistic species (E > 0)

The directionality principles predict that

(i) in the case of large populations, E and S will be negatively correlated, and

(ii) when population size is small, there will exist no correlation between E and S.
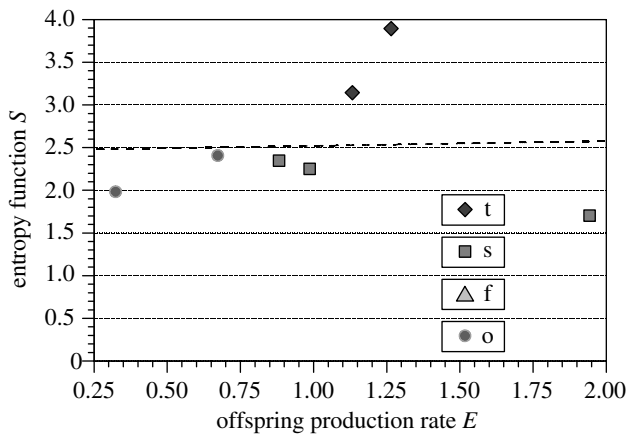
Figure 4. The relation between E and S for species that satisfy the condition $E>0$ (opportunistic species). (Letter codes as in figure 2.)

Table 2. Variation in r for all species with $E<0$, for phylogenetic groups containing species with $E<0$ (o, herbs from disturbed habitats; f, understorey forest herbs; t, trees; s, shrubs), and for all species with $E>0$.
(n, number of species.)

	mean	s.d.	variance	s.e.m.	N
$E<0$	0.037	0.191	0.037	0.025	59
f	0.072	0.172	0.029	0.046	14
o	0.033	0.261	0.068	0.052	25
s	0.017	0.081	0.007	0.033	6
t	0.021	0.062	0.004	0.017	14
$E>0$	0.452	0.283	0.080	0.107	7

The sample size of the opportunistic species is small (six members), a condition we will discuss in interpreting the statistical patterns. Opportunistic species are defined by recurrent episodes of rapid exponential growth when resources are abundant, and periods of population decline, leading to a small population size, when resources become exhausted. The opportunistic species considered are in the early phase of rapid exponential growth. The population size can therefore be assumed as comparatively small. In view of the small population size, we predict that there will exist either (i) no correlation between E and S, or (ii) E and S will be positively correlated.

The correlation between E and S for the sample of opportunistic species is shown in figure 4. The values E and S are not negatively correlated, which is consistent with the prediction.

(iii) *Correlation pattern: implications*

The empirical data indicate that correlation patterns between E and S are as follows: (i) $E<0$: E and S are negatively correlated; and (ii) $E>0$: E and S are not negatively correlated.

In order to evaluate the implications of the correlation pattern between E and S as regards the empirical support for directionality theory, several issues of statistical methodology need to be considered. We first note that the parameters E, S and r are related by the algebraic identity

$$rT = E + S.$$

This equation holds for each species in the group irrespective of the sign of E. Hence, it is valid for both equilibrium and opportunistic species. In view of this identity, we can infer that in both equilibrium and opportunistic species, E and S will be negatively correlated provided that (i) the mean value of r is zero; and (ii) the mean value of r is positive, and the variance in r is sufficiently small.

These observations suggest that the negative correlation observed in the study of equilibrium species may simply be a consequence of the algebraic identity, and thus not reflect the action of evolutionary processes.

In order to assess this possibility and the empirical validity of the theory, we computed the mean and variance of the growth rate for each of the distinct groups in the class of equilibrium species ($E<0$), and the complete set of opportunistic species ($E>0$). Since the opportunistic

species consist only of six members, we made no distinction between the subgroups.

The values of the mean and variance of the growth rate and the sample size of the different groups are given in table 2.

The following relations between the value of E and the correlation between E and S as revealed by the data need to be considered.

(i) $E<0$: the variance of r ranges from 0.004 to 0.068. E and S are negatively correlated. However, there is no correlation between the variance in r and the correlation coefficient between E and S for the groups and subgroups. Shrubs (s) and trees (t) have a comparable low variance in r. However, the correlation coefficient between E and S is small for shrubs and large for trees.
(ii) $E>0$: the variance in r is 0.080, the correlation coefficient between E and S is positive.

We note that E and S are positively correlated when $E>0$, and negatively correlated when $E<0$. We also observe that no correlation exists between the variance in r and the correlation coefficient between E and S. These observations are inconsistent with the hypothesis that the relation between E and S is determined solely by the algebraic constraints on the demographic variables. These algebraic constraints will necessarily influence the magnitude of the correlation, particularly when the variance in r is approximately zero. Our analysis indicates that the negative correlation between E and S when $E<0$, and the positive correlation between E and S when $E>0$, are due primarily to the evolutionary forces driving changes in entropy. Therefore, we conclude that the empirical observations are consistent with the predictions of the theory.

We should emphasize that owing to the small sample size of the opportunistic species, the positive correlation between E and S observed may be an artefact of sample size. However, our argument on the primacy of evolutionary forces and the secondary effect of algebraic constraints in explaining the statistical patterns does not rest uniquely on the correlations described in the opportunistic species. The argument derives mainly from the fact that there exists no correlation between the variance in r and the correlation coefficient between E and S for the different subgroups considered: namely, shrubs, trees, forest herbs and herbs from open disturbed habitats.

7. CONCLUSION

The empirical studies described here are consistent with the predictions of directionality theory: namely, (i) an increase in entropy in populations subject to bounded growth

constraints, and (ii) a decrease or a random non-directional change in entropy in populations subject to unbounded growth constraints. The distribution of the values inherent in these relationships is consistent with the ecological conditions under which different life forms are found. Thus, herbaceous plants from open, disturbed habitats (low entropy species) and trees (high entropy species) occupy the extremes predicted by the theory, with herbs from relatively stable habitats and shrubs occupying intermediate positions. It would be simplistic, however, to assume that evolution (in the long term) and succession (in the short term) proceed from short-lived herbs to long-lived trees when conditions are stable, and in the opposite direction when disturbance dominates. The conditions leading to the dominance of woody versus non-woody species may be determined by the precise combination of resources for which plants compete (e.g. nutrients versus light) as well as the intensity and frequency of disturbance (Tilman 1988). Other complications, such as the evolution of multiple modes of reproduction, the decoupling of the seed and vegetative stages and the existence of conditions leading to semelparity, are elements that the theory does not address in its present form. The main claims, however, that entropy increases in equilibrium species, and decreases or undergoes random non-directional changes in opportunistic species, are supported by the cross-species comparisons we have documented.

REFERENCES

Austad, S. N. 1997 Comparative aging and life histories in mammals. *Exp. Gerontol.* **32**, 23–38.

Billingsley, P. 1965 *Ergodic theory and information.* New York: Wiley.

Bonner, J. T. 1988 *The evolution of complexity.* Princeton University Press.

Calder, W. 1996 *Size, function and life history.* New York: Dover Publications.

Charlesworth, B. 1994 *Evolution in age-structured populations.* Cambridge University Press.

Charlesworth, B. & Williamson, J. A. 1975 The probability of survival of a mutant gene in an age-structured population and implications for the evolution of life histories. *Genet. Res.* **26**, 3–10.

Coale, A. I. 1972 *The growth and structure of human populations. A mathematical investigation.* Princeton University Press.

Cochran, M. E. & Ellner, S. 1992 Simple methods for calculating age-based life history parameters for stage-structured populations. *Ecol. Monogr.* **62**, 345–364.

Crow, J. F. 1979 Gene frequency and fitness change in an age-structured population. *Ann. Hum. Genet.* **42**, 355–372.

Crow, J. F. & Kimura, M. 1970 *An introduction to population genetics theory.* New York: Harper & Row.

Demetrius, L. 1974 Demographic parameters and natural selection. *Proc. Natl Acad. Sci. USA* **71**, 4645–4647.

Demetrius, L. 1977 Measures of fitness and demographic stability. *Proc. Natl Acad. Sci. USA* **74**, 384–388.

Demetrius, L. 1983 Statistical mechanics and population biology. *J. Stat. Phys.* **30**, 709–753.

Demetrius, L. 2000 Directionality theory and the evolution of body size. *Proc. R. Soc. B* **267**, 2385–2391. (doi:10.1098/rspb.2000.1295.)

Demetrius, L. 2001 Mortality plateaus and directionality theory. *Proc. R. Soc. B* **268**, 2029–2037.

Demetrius, L. 2004 Caloric restriction, life span and entropy. *J. Gerontol. Biol. Sci.* **59A**, 902–935.

Demetrius, L. & Gundlach, V. M. 1999 Evolutionary dynamics in random environments. In *Stochastic dynamics* (ed. H. Crauel & V. M. Gundlach), pp. 371–394. New York: Springer.

Demetrius, L. & Gundlach, V. M. 2000 Game theory and evolution: finite size and absolute fitness measures. *Math. Biosci.* **168**, 9–38.

Demetrius, L. & Ziehe, M. 1984 The measurement of Darwinian fitness in human populations. *Proc. R. Soc. B* **222**, 33–50.

Demetrius, L., Gundlach, V. M. & Ochs, G. 2004 Complexity and demographic stability in population models. *Theor. Popul. Biol.* **65**, 211–225.

Emlen, J. M. 1984 *Population biology: the coevolution of population dynamics and behavior.* New York: McMillan Publishing Co.

Fisher, R. A. 1958 *The genetical theory of natural selection*, 2nd edn. New York: Dover Publications Inc.

Franco, M. & Silvertown, J. 1996 Life history variation in plants: an exploration of the fast–slow continuum hypothesis. *Phil. Trans. R. Soc. B* **351**, 1341–1348.

Harvey, P. H. & Pagel, M. D. 1991 *The comparative method in evolutionary biology.* Oxford University Press.

Keyfitz, N. 1972 Population waves. In *Population dynamics* (ed. T. N. E. Greville), pp. 1–39. New York: Academic Press.

Kim, Y. J. & Schoen, R. 1993 On the intrinsic force of convergence to stability. *Math. Popul. Stud.* **4**, 89–102.

Kowald, A. & Demetrius, L. 2005 Directionality theory: a computational study of an entropic principle in evolutionary processes. *Proc. R. Soc. B* **272**, 741–749. (doi:10.1098/rspb.2004.3012.)

Lande, R. 1982 A quantitative genetic theory of life history evolution. *Ecology* **3**, 607–617.

Lawton, J. & Brown, T. 1986 The population and community ecology of invading insects. *Phil. Trans. R. Soc. B* **314**, 607–614.

MacArthur, R. H. & Wilson, E. O. 1967 *The theory of island biogeography. Monographs in Population Biology*, vol. 1. Princeton University Press.

Pianka, E. R. 1970 On *r*- and *K*-selection. *Am. Nat.* **104**, 592–597.

Pollak, E. 1976 A stochastic treatment of rare genes in large populations with overlapping generations. *Theor. Popul. Biol.* **10**, 109–126.

Reznick, D., Bryant, M. J. & Bashey, F. 2002 *r*- and *K*-selection revisited: the role of population regulation in life-history evolution. *Ecology* **83**, 1509–1520.

Sinai, Ya. G. 1989 Kolmogorov's work on ergodic theory. *Ann. Prob.* **17**, 833–839.

Stearns, S. C. 1992 *The evolution of life histories.* Oxford University Press.

Tilman, D. 1988 *Plant strategies and the dynamics and structure of plant communities. Monographs in Population Biology*, vol. 26. Princeton University Press.

Tuljapurkar, S. 1982 Why use population entropy? It determines the rate of convergence. *J. Math. Biol.* **13**, 225–237.