JOURNAL OF Evolutionary Biology

doi: 10.1111/j.1420-9101.2011.02378.x

REVIEW

Natural selection. I. Variable environments and uncertain returns on investment*

S. A. FRANK

Department of Ecology and Evolutionary Biology, University of California, Irvine, CA, USA

Keywords:

economics;
mathematical theory;
population genetics.

Abstract

Many studies have analysed how variability in reproductive success affects fitness. However, each study tends to focus on a particular problem, leaving unclear the overall structure of variability in populations. This fractured conceptual framework often causes particular applications to be incomplete or improperly analysed. In this article, I present a concise introduction to the two key aspects of the theory. First, all measures of fitness ultimately arise from the relative comparison of the reproductive success of individuals or genotypes with the average reproductive success in the population. That relative measure creates a diminishing relation between reproductive success and fitness. Diminishing returns reduce fitness in proportion to variability in reproductive success. The relative measurement of success also induces a frequency dependence that favours rare types. Second, variability in populations has a hierarchical structure. Variable success in different traits of an individual affects that individual's variation in reproduction. Correlation between different individuals' reproduction affects variation in the aggregate success of particular alleles across the population. One must consider the hierarchical structure of variability in relation to different consequences of temporal, spatial and developmental variability. Although a complete analysis of variability has many separate parts, this simple framework allows one to see the structure of the whole and to place particular problems in their proper relation to the general theory. The biological understanding of relative success and the hierarchical structure of variability in populations may also contribute to a deeper economic theory of returns under uncertainty.

Introduction

Natural selection favours traits that enhance fitness. But how does one measure fitness? Several studies have shown that it is not just average reproductive success that matters. Variation in reproductive success also plays an important role in determining long-term evolutionary trends. To understand the basic notions of fitness and evolutionary change by natural selection, one must understand the particular consequences of different kinds of variation.

The literature on variation splits into two groups. On the one side, bits of folk wisdom dominate thinking. The slogan that natural selection maximizes geometric mean fitness is one example. Such folk wisdom is true in special cases. But as a guiding principle, the simple geometric mean slogan misleads as often as it helps.

On the other side, a technically demanding specialist literature divides into numerous distinct ways of framing the problem. Each technical expression emphasizes a particular aspect of variation, refining unique examples at the expense of providing a coherent view of the whole.

This article provides a tutorial on the different kinds of variation and their evolutionary consequences. I emphasize simple examples to develop understanding of temporal, spatial, developmental and trait variation. Each type of variation was originally studied as a separate problem. In this tutorial, I follow Frank & Slatkin (1990), who showed that these seemingly different types of variation can be understood in a unified way. The unified framework arises from two steps.

Correspondence: Steven A. Frank, Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697–2525, USA.
Tel.: +1 949 824 2244; fax: +1 949 824 2181; e-mail: safrank@uci.edu
*Part of the Topics in Natural Selection series (see Box 1).

Box 1: Topics in the theory of natural selection

This article is the first in a series on natural selection. Although the theory of natural selection is simple, it remains endlessly contentious and difficult to apply. My goal is to make more accessible the concepts that are so important, yet either mostly unknown or widely misunderstood. I write in a nontechnical style, showing the key equations and results rather than providing full derivations or discussions of mathematical problems. Boxes list technical issues and brief summaries of the literature.

First, it is relative reproduction that matters. Only those traits associated with relatively greater success than average increase over time. Relative measures of success induce diminishing returns: a doubling of reproduction provides less than a doubling of relative success (Gillespie, 1977; Frank & Slatkin, 1990). With diminishing returns, increasing variation in reproductive success reduces fitness.

Second, the different types of variation can be expressed as different levels within a unified hierarchy (Frank & Slatkin, 1990). Variable success in different traits of an individual affects that individual's variation in reproduction. Correlation between different individuals' reproduction affects variation in the aggregate success of particular alleles across the population. Temporal, spatial and developmental variation affect the way in which individual variations combine to determine the overall variability in the number of copies produced by a particular allele.

I also discuss the relation of economic theories of risk and uncertainty to evolutionary theories of variability.

Relative success induces diminishing returns

The success of genes and of traits must ultimately be measured by their relative frequency in a population. The calculation of relative frequency leads to surprising consequences when there is variability (Gillespie, 1977; Frank & Slatkin, 1990; Orr, 2007).

To illustrate the problem, consider two alternative types in a population, A_1 and A_2. I will often refer to the alternative types as alleles at a genetic locus. However, the same analysis would apply to any heritable alternative types in a population that have the same essential properties as alleles.

Some simple notation helps to express the argument. Each definition uses subscripts to associate with the alternative alleles, A_1 and A_2, respectively. Define q_1 and q_2 as the allele frequencies, such that $q_1 + q_2 = 1$. Let R_1 and R_2 measure reproductive success, the average number of descendant copies produced by each parental allele. The average reproductive success in the population is $\bar{R} = q_1R_1 + q_2R_2$. The success of individual parental copies has a random component. Thus, all of the measures of reproductive success fluctuate randomly. Throughout, the unqualified words *average* and *mean* refer to the arithmetic average.

The frequency of A_1 after one round of reproduction is

$$q_1' = q_1(R_1/\bar{R}) = q_1F_1, \tag{1}$$

where F measures relative success. I use the word *fitness* for relative success. This equation shows that fitness ultimately determines the success of an allele. Reproductive success, R, influences fitness. But the key relationship between reproductive success and fitness is mediated through the definition for fitness

$$F_1 = R_1/\bar{R} = \frac{R_1}{q_1R_1 + q_2R_2}. \tag{2}$$

Figure 1 illustrates the two key properties of fitness. First, fitness increases at a diminishing rate with a rise in reproductive success (Gillespie, 1977; Frank & Slatkin, 1990). Put another way, the fact that fitness is a relative measure means that linear changes in reproductive success translate into nonlinear changes in fitness.

Second, the curvature of the relation between reproductive success and fitness is frequency dependent (Frank & Slatkin, 1990). A rare type has a nearly linear relation between reproduction and fitness. A common type has a very strongly diminishing relation between reproduction and fitness. This means that rare and common types are influenced differently by the consequences of variability, because more strongly diminishing returns cause variability in reproductive success to impose a greater penalty on fitness.

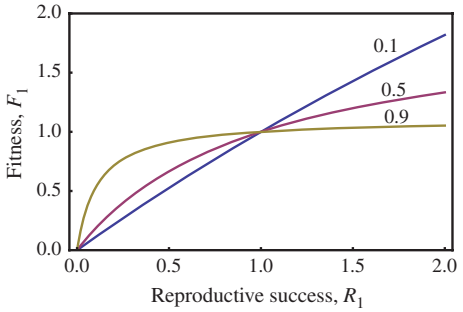

Fig. 1 The curvature of fitness vs. reproductive success depends on allele frequency. The plots here illustrate eqn (2). The numbers above each curve show q_1, the frequency of the allele for which the relationship is plotted. If $q_1 = 0.1$, the relationship follows the upper curve for allele A_1; the lower curve can then be interpreted as the relationship for allele A_2 with frequency $q_2 = 0.9$. The difference between the upper and lower curves illustrates the frequency dependence of the relation between fitness and reproductive success. Note that there is little curvature when an allele is rare, which leads to an advantage for rare types. Redrawn from Frank & Slatkin (1990).

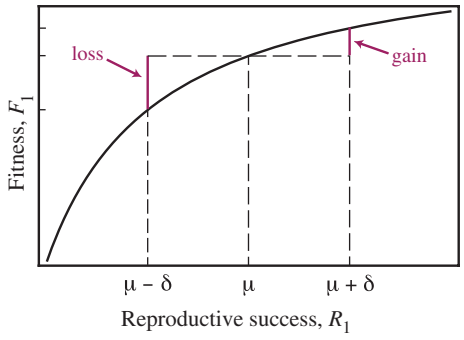

Fig. 2 Increasing variation in reproductive success reduces fitness. Expected reproductive success is μ. Fluctuations of $\pm\ \delta$ occur. Positive fluctuations return a smaller gain in fitness than the loss suffered from a negative fluctuation. Thus, equally frequent positive and negative fluctuations return a net loss. Redrawn from Frank & Slatkin (1990).

Figure 2 shows that diminishing returns cause a loss of fitness. In the figure, expected reproductive success, R, is μ. Deviations of $\pm\ \delta$ occur, with increases and decreases at equal frequencies. The gain in relative fitness, F, for an increase of δ units of reproductive success is less than the corresponding loss in fitness when reproductive success is reduced by δ. Expected fitness therefore declines as the frequency and magnitude of deviations increase. Note that the discount to fitness rises as the curvature between fitness and reproductive success increases.

Reproduction multiplies and variation reduces success

Suppose an individual has two offspring. Each of those offspring has two offspring. The original individual has four grandchildren. Compare that output to a second individual that has three offspring, and each of those offspring has one offspring. The second individual has three grandchildren. In each case, the average reproduction per generation is two offspring. However, the individual with less variable reproduction has greater success than the individual with more variable reproduction.

The difference occurs because reproduction multiplies over time. The value $2 \times 2 = 4$ is greater than $3 \times 1 = 3$, even though the arithmetic averages are the same in each case. In multiplicative series, variation reduces the multiplicative product. Rather than measuring success by the arithmetic average, such as $(3 + 1)/2 = 2$, the proper average in a multiplicative series is the number that, when used to multiply in each generation, gives the total output. In the reproductive series over two generations of $3 \times 1 = 3$, we need a number that when multiplied by itself gives three. This works out as $\sqrt{3} \times \sqrt{3} = 3$, so the multiplicative mean is $\sqrt{3} \approx 1.73$. The multiplicative mean is usually called the *geometric mean*.

Approximation for the geometric mean

A simple approximation of the geometric mean is often useful. Suppose the arithmetic mean of a series is μ, and the variance of the series is σ^2. Then, the geometric mean is approximately $\mu - \sigma^2/2\mu$. This approximation shows clearly how the variance reduces the geometric mean.

For example, in the series $3 \times 1 = 3$, the mean is 2, the variance is 1, the true geometric mean is 1.73, and the approximation gives the geometric mean as 1.75. The smaller the variance is in relation to the mean, the better the approximation. In evolutionary models, one often makes the assumption that average reproductive success is close to one, $\mu \approx 1$, so that the approximation for the geometric mean often appears as $\mu - \sigma^2/2$.

The geometric mean principle

Reproduction multiplies. The greatest multiplicative series has the highest geometric mean reproduction. Thus, the type with greatest long-term fitness would appear to be the type with the highest geometric mean reproductive success. That conclusion is often called the geometric mean principle.

The highest geometric mean is sometimes associated with greatest evolutionary success. However, geometric mean reproductive success by itself often misleads, because it hides more than it helps. The problem of evolutionary success and fitness turns out to be more subtle and more interesting. The following sections explain why.

Population regulation and relative success

The total resources available to the population limit reproductive success. That density-dependent competition causes the reproductive success of each type to be influenced by the reproduction of other types. For that reason, one cannot simply multiply the reproductive successes of each type independently and then compare the long-term geometric means. Instead, each bout of density-dependent competition causes interactions between the competing types. Those interactions depend on frequency. Reproduction of a rare type has little competitive effect on a common type. Reproduction of a common type has a strong competitive effect on a rare type (Fig. 1).

The fitness measure of relative success in eqn (2) accounts for density-dependent interactions. The particular way in which density-dependent competition arises has important consequences.

Expected change in frequency

We can circumscribe the main conceptual issues by focusing on the expected (average) change in allele

frequency. Each process can be studied with respect to whether it tends to increase or decrease the expected frequency change.

Equation (1) gives the definition for the allele frequency in the following generation, $q_1' = q_1 F_1$. The change in allele frequency is $\Delta q_1 = q_1' - q_1 = q_1(F_1 - 1)$. Using the definition for fitness in eqn (2) allows us to write the change in frequency as

$$\Delta q_1 = q_1 q_2 \left(\frac{R_1 - R_2}{\bar{R}} \right). \tag{3}$$

The reproductive successes fluctuate randomly. Because those random fluctuations occur in both the numerator and the denominator, there is no simple way to express the exact change in frequency. If we assume that the fluctuations in success are small relative to the average success, and we normalize all of the successes so that they are close to one, then we can write the approximate expected change in frequency as

$$\mathrm{E}(\Delta q_1) \approx q_1 q_2 \{ (\mu_1 - \mu_2) + [\mathrm{cov}(R_2, \bar{R}) - \mathrm{cov}(R_1, \bar{R})] \}, \tag{4}$$

where μ_1 and μ_2 are the expected reproductive successes for types 1 and 2, respectively. This equation, from Frank & Slatkin (1990), is equivalent to an approximation given by Gillespie (1977).

The expected change is the average tendency. Because of the inherent fluctuations in success, the actual change in frequency in any generation may be in the opposite direction of the expected change. Over long time periods, three different patterns of evolutionary dynamics occur.

First, if the random fluctuations in the average reproductive successes of the types are large relative to the directional bias in eqn (4), then randomness dominates. The type frequencies will bounce up and down in a nearly neutral way. Eventually, one type will become fixed, and the other will disappear from the population. If we start at frequency q_1, then the probability that type A_1 fixes is q_1, and the probability that type A_2 fixes is $q_2 = 1 - q_1$. True fixation only occurs in finite populations. In infinite populations, the related notion of quasi-fixation arises as the frequency of a type becomes very close to one. To keep things simple, I ignore important technical distinctions between finite and infinite populations (Gillespie, 1994; Ewens, 2010).

Second, if the directional bias is much larger than the fluctuations in the average reproductive successes of the types, then the type frequencies change in an almost deterministic way. If the direction of change remains the same across the range of type frequencies, then the type with the greater expected success will usually become fixed in a relatively short period of time. If frequency dependence provides a sufficient advantage to the rare type, then the direction of change may shift with type frequency in a way that tends to maintain both types in the population.

Third, if random fluctuations are of roughly the same magnitude as the directional bias, then frequency changes combine both directional and random aspects. In some cases, frequencies will fluctuate, and both types will remain in the population for a very long time. In other cases, frequencies may fluctuate for a significant period of time, but eventually one type or the other will become fixed. Fixation will be biased towards the type favoured by the directional tendency set by eqn (4). However, the other type may occasionally fix because of the random fluctuations.

I emphasize the major processes that influence the directional tendency in eqn (4). In particular, the hierarchical structure of variability sets the directional tendency, which in turn shapes the qualitative patterns of dynamics. I make only brief comments on long-term evolutionary dynamics, which require technical analysis of mathematical models to evaluate fully (Gillespie, 1994; Ewens, 2010). As noted in the previous paragraphs, long-term dynamics include issues such as the probability that a particular genotype becomes fixed and the maintenance of polymorphism.

Hierarchical structure of variability

Populations have a naturally nested hierarchical structure when considering genetics. Populations are composed of genotypes, and genotypes are composed of individuals. In a haploid model with two alternative alleles, there are two genotypes in the population, and numerous copies (individuals) of each allelic type. The hierarchical structure of variability makes explicit the variances and the correlations at different levels in the hierarchy.

For example, the reproductive success R_1 is the average success taken over all copies of the allele A_1. Similarly, R_2 is the average over all copies of A_2. To analyse the variability in the aggregate success of allele A_1, one must consider the variability in the success of each copy of A_1 and the correlations in success between different copies. The analysis for A_2 must consider the variability in the success of each allelic copy and the correlations in success between the alleles.

We can relate the variances of individual reproductive success to the variances and covariances for the different allelic types by writing

$$\mathrm{var}(R_1) = \rho_1 \sigma_1^2 \tag{5a}$$

$$\mathrm{var}(R_2) = \rho_2 \sigma_2^2 \tag{5b}$$

$$\mathrm{cov}(R_1, R_2) = \rho_{12} \sigma_1 \sigma_2 \tag{5c}$$

where ρ_1, ρ_2 and ρ_{12} are the correlations in reproductive success between randomly chosen pairs of A_1, A_2 or A_1 and A_2 individuals, respectively.

Frank & Slatkin (1990) introduced this explicit partitioning of variances and covariances for types into their individual components. Any realistic analysis of variability must make explicit the individual-level fluctuations and the associations between individuals. Although this explicit treatment of variability is fundamental, the partitioning of variances and covariances in eqn (5) has often been regarded as some sort of highly technical or specialized analysis. This mistake has limited progress in understanding fitness with respect to spatial and temporal fluctuations in success.

If we combine eqns (4) and (5) and keep things simple by assuming that the correlation between types is zero, $\rho_{12} = 0$, we obtain

$$\mathrm{E}(\Delta q_1) \approx q_1 q_2 \left\{ (\mu_1 - q_1 \rho_1 \sigma_1^2) - (\mu_2 - q_2 \rho_2 \sigma_2^2) \right\}, \qquad (6)$$

which means that, on average, type A_1 increases in frequency when

$$\mu_1 - q_1 \rho_1 \sigma_1^2 > \mu_2 - q_2 \rho_2 \sigma_2^2. \qquad (7)$$

The following sections show that different kinds of variability can be understood by the hierarchical partitioning of associations between traits within an individual and associations between different individuals. Variability interacts with the processes of density-dependent population regulation.

Temporal variability

Dempster (1955) introduced a model of temporal variation in which all alleles of the same type have the identical reproductive success within a generation, $\rho_1 = \rho_2 = 1$, and there is no correlation between types, $\rho_{12} = 0$. In this haploid model, each individual has one allele, either A_1 or A_2. The condition for the expected increase of type 1 from eqn (7) is

$$\mu_1 - q_1 \sigma_1^2 > \mu_2 - q_2 \sigma_2^2. \qquad (8)$$

This equation illustrates the rare-type advantage induced by density-dependent population regulation. When the frequencies are equal, $q_1 = q_2 = 1/2$, then the condition favours the type with the higher geometric mean fitness, $\mu_1 - \sigma_1^2/2$ vs. $\mu_2 - \sigma_2^2/2$.

As the frequencies approach one boundary, $q_1 \to 0$ and $q_2 \to 1$, the condition to favour A_1 becomes $\mu_1 > \mu_2 - \sigma_2^2$. At the other boundary, the condition favouring A_1 becomes $\mu_1 - \sigma_1^2 > \mu_2$. Thus, the directional tendency often shifts with frequency.

In spite of the inherent rare-type advantage, polymorphism is not maintained in this haploid model (Gillespie, 1973; Hartl & Cook, 1973; Karlin & Lieberman, 1974). The high variance in fluctuations eventually causes one of the types to fix (or to become nearly fixed in an infinite population). The type with the higher geometric mean success has the advantage at the frequency midpoint. That type fixes with higher probability. If the geometric means for the two types are close to each other, then frequencies may fluctuate for a long time, and the bias towards fixing the favoured type is relatively weak. If the geometric means for the two types are significantly different, then fixation happens sooner and with a stronger bias towards the favoured type.

Correlations and genotypic homoeostasis

In the previous section, I assumed that all individuals carrying the same allele have the same reproductive success in each generation. In that case, all of the variation arises from the response of an allele to environmental fluctuations, with no variation between individuals of the same genotype. No variation means that individuals of the same type are perfectly correlated, $\rho_1 = \rho_2 = 1$.

Alternatively, different individuals of the same type may respond differently to environmental fluctuations. There are many ways to express individual variation. For example, individual responses may fluctuate about the long-term arithmetic mean, μ, and the pairwise correlation between individuals in each generation may be ρ (Frank & Slatkin, 1990). In that case, the variance in the average reproductive success of A_1 is $\rho_1 \sigma^2$, with a similar expression for A_2.

Reduced correlation between individuals lowers the variation in the average success of a type. That relation arises from the fact that the variance of an average is reduced by the number of uncorrelated observations in the sample. We can express the effective sample size of uncorrelated observations as $n^* = 1/\rho$, so that the variance of the mean, σ^2/n^*, is $\rho\sigma^2$.

One can think of the pairwise correlations between individuals of the same genotype as the *genotypic homoeostasis*. If all individuals of a genotype respond in exactly the same way to each environmental state, then the correlation between pairs of individuals is perfect, $\rho = 1$. That perfect correlation increases the variance in the average reproductive success of the genotype. One can think of such strong correlation as strong homoeostasis or canalization of development for the genotype. By contrast, weak correlation between individuals of the same genotype, with low values of ρ, corresponds to greater developmental fluctuations and relatively weaker genotypic homoeostasis.

Given the variances in the average reproductive successes of the types as $\rho\sigma^2$, the condition for A_1 to be favoured is given in eqn (7). In a haploid model, the long-term bias in fixation depends on the relative geometric means derived when frequencies are equal, $q_1 = q_2 = 1/2$, yielding the condition for A_1 to be favoured as

$$\mu_1 - \rho_1 \sigma_1^2/2 > \mu_2 - \rho_2 \sigma_2^2/2. \qquad (9)$$

This expression shows that temporal fluctuations favour reduced genotypic homoeostasis, with low values

Box 2: Optimal phenotypes in response to environmental variability

Individuals may be able to match their phenotype to particular environments. Phenotypic plasticity occurs when an organism can sense the particular environmental state and adjust its traits accordingly. If organisms do not adjust their phenotypes in response to the particular environmental state, then they may produce a stochastic response tuned to the pattern of fluctuation. A stochastic response is sometimes called *bet hedging*.

Bet hedging can increase the aggregate success of a genotype or strategy. Suppose, for example, that the environment is equally likely to be in one of two states. For each state, there is a different optimal phenotypic response. However, the organism cannot adjust its phenotype in response to the particular state. If each individual of the genotype has a random component to its phenotypic expression, then in each generation, some individuals will match the environment with the best phenotype and some will not. The mixture of phenotypic expressions reduces the variance of the aggregate success of the genotype by reducing the correlation between individuals of that genotype.

The concept of reduced correlation between individuals of a genotype highlights an essential aspect of the bet hedging problem. Frank & Slatkin (1990) emphasized the general point about correlations between members of a genotype, as shown in eqn (5). McNamara (1995) independently described a similar interpretation, but referred to the process as kin selection, a label that I would avoid in this case. Correlations between relatives do matter, but not in the way that one usually associates with the costs and benefits of social behaviour in kin selection.

Bet hedging can also arise within an individual. For example, the individual's alternative traits may be expressed stochastically in separate bouts of resource acquisition, in a way that reduces the overall variance in success. Reduced variance in resource acquisition typically provides increased expected reproductive success, because success rises in a diminishing way with resources.

To get started on the literature, here are a few recent overviews for phenotypic plasticity (Stearns, 1989; Houston & McNamara, 1992; Moran, 1992; Scheiner, 1993; Via *et al.*, 1995; Pigliucci, 2001; West-Eberhard, 2003; DeWitt & Scheiner, 2004) and bet hedging (Cooper & Kaplan, 1982; Seger & Brockmann, 1987; Sasaki & Ellner, 1995; Grafen, 1999; Wilbur & Rudolf, 2006; King & Masel, 2007; Donaldson-Matasci *et al.*, 2008). An interesting information theory approach may provide a useful connection between these subjects (Kussell & Leibler, 2005).

All of these cases analyse phenotypic adjustment or phenotypic stochasticity. In these cases, one must also account for the diminishing relation between reproductive success and fitness (Fig. 1) and the ways in which the correlations between individuals determine the mean and variance of aggregate success for each type (eqn 5).

of ρ. A type with low average reproductive success, μ, can be favoured if it also has low genotypic homoeostasis, ρ, reducing its variance in average reproductive success sufficiently to give it a higher geometric mean fitness than its competitor. Reduced genotypic homoeostasis is a general expression of the widely discussed problem of bet hedging (see Box 2).

Developmental variability

Gillespie (1974a) introduced a model in which the reproductive success of each of the N haploid individuals in the population depends on its interactions with the environment during development. The reproductive successes of different individuals are independent because, by Gillespie's assumptions, different individuals experience different conditions and develop in an uncorrelated way. Nevertheless, the finite population size ensures that an individual's reproductive success correlates with the average reproductive success of its genotype.

The correlation of two randomly chosen A_1 alleles is $\rho_1 = 1/(Nq_1)$, because there are Nq_1 individuals of type A_1, and hence a chance $1/(Nq_1)$ of choosing the same individual twice. By the same reasoning, $\rho_2 = 1/(Nq_2)$. The correlation between types is zero, because different individuals experience different conditions. Substituting these values into eqn (8), we find that A_1 increases for any allele frequency when

$$\mu_1 - \sigma_1^2/N > \mu_2 - \sigma_2^2/N.$$

Because this condition no longer depends on allele frequencies, it is sufficient to describe long-term evolutionary advantage without the need to consider frequency dependence. An allele with a long-term advantage is more likely to become fixed than a neutral allele with the same initial gene frequency.

Gillespie (1974a) presented this model of individual developmental variation as a separate problem from the general analysis of fluctuations in reproductive success. The analysis here, from Frank & Slatkin (1990), shows that individual variation is just a special case of the general model of temporal variation. One obtains the case of individual variation by properly calculating the correlations in reproductive success between individuals.

Spatial variability and local population regulation

The classical Dempster (1955) model of temporal variation assumes that density-dependent regulation occurs in one large population. In that model, density regulation induces frequency dependence that favours the rare genotype. I mentioned earlier that, in spite of the rare-type advantage, one of the types eventually becomes fixed, because the random fluctuations in frequency are

too strong relative to the directional tendency of evolutionary change. Fixation is biased towards the type with the highest geometric mean.

In a different model, Levene (1953) showed that spatial variation does maintain genetic polymorphism. In the Levene model, there are many independent spatial locations. Each location has its own independent density-dependent competition for resources.

Gillespie (1974b, 1978) showed that one can think of the Levene model of spatial variation as the sum of K independent models of temporal variation. If there is only one patch, $K = 1$, then reproductive successes fluctuate over time in that patch, and all competition occurs in that single patch. This model is identical to the classical Dempster model for temporal variation.

As K increases, each independent patch fluctuates with the same rare-type advantage of the classical Dempster model. The total fluctuation in each generation is the average of the fluctuations over all patches. Because the patches fluctuate independently, the variance of the average fluctuation over the entire population is reduced by $1/K$. This reduction arises because the variance of the mean for K independent observations is the variance of each observation divided by K.

Gillespie (1974b, 1978) provided the full analysis for this averaging over K patches. However, he did not provide a simple interpretation or a simple expression for how fluctuations lead to a particular level of polymorphism when the number of patches is large.

Frank & Slatkin (1990) noted that, as K becomes large, the population-wide fluctuations in each generation become small because of the averaging effect over the many patches. Thus, we can treat eqn (6) as an essentially deterministic process. The rare-type frequency dependence now dominates. The equilibrium frequency of types can be obtained by solving $\mathrm{E}(\Delta q_1) = 0$, which yields

$$\frac{q_1}{q_2} = \frac{\mu_1 - \mu_2 + \rho_2\sigma_2^2}{\mu_2 - \mu_1 + \rho_1\sigma_1^2}$$

as given by Frank & Slatkin (1990). Here, each ρ is the correlation between copies of an allelic type measured within each patch. This result shows that naive comparison of geometric mean success is not sufficient to understand evolutionary outcome.

Trait variability within individuals

The theory above takes an individual's average and variance in reproductive success as given parameters. However, the actual processes that lead to individual means and variances arise from the way that individuals acquire resources and produce offspring, including acquisition of food, protection from predators and so on. To analyse the full hierarchy of causes for variability, we should begin with the question: How does the allocation of an individual's resources among alternative traits influence that individual's mean and variance in reproductive success?

I use the example of traits for resource acquisition. The same analysis applies to any trait that influences reproductive success, such as defence against parasites.

One trait

Let us start with a single trait. The return of resources on investment has a random component, δ. The random component of resource acquisition affects reproductive success by an amount $f(\delta)$. Then, a simple way to write the reproductive success is

$$R = 1 + f(\delta).$$

If we assume that the random fluctuations, δ, have a mean of zero and a variance of V_x, and that the fluctuations are relatively small, then the average reproductive success is approximately

$$\mu \approx 1 + f''V_x/2,$$

where f'' is the second derivative of f evaluated at zero (Real, 1980; Stephens & Krebs, 1986). Typically, one assumes that fluctuations in traits for resource acquisition have a diminishing return shape as in Fig. 2, in which case $f'' < 0$. Thus, greater fluctuations, V_x, reduce expected reproductive success. All else equal, resource acquisition strategies with less variability yield higher average reproductive success than those strategies with more variability.

The variance in an individual's reproductive success is approximately

$$\sigma^2 \approx \mathrm{var}(f'\delta) = f'^2V_x, \tag{10}$$

where f' is the first derivative of f evaluated at zero.

A full evaluation of fitness requires specifying the means, variances and correlations between all individuals in the population. The correlations must be evaluated in relation to the heritable types we are following over time. The earlier sections provided the methods for studying evolutionary dynamics in relation to fitness.

Here, to keep the focus on trait variability within individuals, I give only the geometric mean reproductive success for an individual. The geometric mean for each individual accounts for the average and variance in individual reproductive success but neither the correlations between types nor the role of density dependence in fitness. Assuming that the fluctuations in returns, V_x, are relatively small, the geometric mean reproductive success is approximately

$$G = \mu - \sigma^2/2\mu \approx 1 + \left(f'' - f'^2\right)V_x/2. \tag{11}$$

Two traits

How should an individual invest in two different traits that provide additive returns? Let reproductive success be

Box 3: Trait variability with two traits

This box shows the details that lead from eqns (12–13). Starting with eqn (12) for R, the average reproductive success is approximately

$$\mu \approx 1 - y\gamma + xf''V_x/2 + yg''V_y/2.$$

The variance in success is approximately

$$\begin{aligned}\sigma^2 &\approx \mathrm{var}(xf'\delta + yg'\epsilon)\\ &= x^2f'^2V_x + y^2g'^2V_y.\end{aligned}$$

The geometric mean is approximately

$$\mu - \sigma^2/2\mu \approx 1 - y\gamma + x\left(f'' - xf'^2\right)V_x/2 + y\left(g'' - yg'^2\right)V_y/2.$$

If we assume that the random component of the two traits is the same, $V_x = V_y$, and $f \equiv g$, then the geometric mean is

$$\begin{aligned}\mu - \sigma^2/2\mu &\approx 1 - y\gamma + f''V_x/2 - (x^2+y^2)f'^2V_x/2\\ &= 1 + \left(f'' - f'^2\right)V_x/2 + f'^2\left[1 - (x^2+y^2)\right]V_x/2 - y\gamma.\end{aligned}$$

The substitutions given in the text lead directly to eqn (13).

$$R = x[1 + f(\delta)] + y[(1-\gamma) + g(\epsilon)] \tag{12}$$

for investment amounts $x + y = 1$, with x and y the fractions of total resources invested in each trait. Here, γ is the discount in expected return for the second trait, and ϵ is the random fluctuation associated with the second trait. The discount, γ, and the fluctuation, ϵ, are small relative to the baseline return of one. The mean of ϵ is zero, and the variance is V_y. I assume that δ and ϵ are uncorrelated. Box 3 lists some of the intermediate steps. Here, I focus on the key result.

If we assume that the random component of each trait is the same, $V_x = V_y$, and $f \equiv g$, then the geometric mean is

$$\mu - \sigma^2/2\mu \approx G + B(x,y), \tag{13}$$

where G is the geometric mean given in eqn (11) for allocating all resources to the first trait, $x = 1$, and $B(x,y)$ is the benefit obtained when mixing allocation of resources between the two traits such that $x + y = 1$, with

$$B(x,y) = f'^2\left[1 - (x^2+y^2)\right]V_x/2 - y\gamma.$$

If we optimize B to obtain the best mixture of allocations between the two traits, we obtain

$$x^* = \frac{1}{2}\left(1 + \frac{\gamma}{\sigma^2}\right) \tag{14a}$$

$$y^* = \frac{1}{2}\left(1 - \frac{\gamma}{\sigma^2}\right), \tag{14b}$$

where γ is the discount in expected return for the second trait given in eqn (12), and σ^2 is the variance in individual reproductive success per trait given in eqn (10).

It pays to invest some resources in the trait with lower expected return as long as $\gamma/\sigma^2 < 1$. The lower expected return is offset by the reduced variance obtained from averaging the returns over two uncorrelated traits. In both biology and financial investing, returns tend to multiply over time. Thus, reduced fluctuations enhance the multiplicative (geometric) average return. In financial investing, the central role of the geometric mean is well known in theory (Bernstein & Wilkinson, 1997), but often ignored in practice (MacBeth, 1995).

An example

The concepts in the previous section are simple. The variance of an average declines with additional uncorrelated components. Reduced variance provides a benefit when success multiplies over time. The technical expressions of those results may obscure the simplicity of the concepts. This section provides a numerical example.

Suppose an organism has two different behaviours by which it can obtain calories. To keep the problem simple, assume that there is a linear relation between calories and reproduction, $f' = 1$. For the first behaviour, the return is on average $\mu_1 = 1.0$ calories, with a variance in return of $\sigma^2 = 0.1$. The second behaviour has a lower average return of $\mu_2 = 1.0 - 0.02 = 0.98$ calories, with the same variance of $\sigma^2 = 0.1$.

If all investment is devoted to the first behaviour, then the geometric mean success is $\mu_1 - \sigma^2/2 = 0.95$. If all investment is devoted to the second behaviour, then the geometric mean success is $\mu_2 - \sigma^2/2 = 0.93$.

In this case, I have assumed $\gamma = 0.02$ and $\sigma^2 = 0.1$. From eqn (14), the optimal allocation to the two traits is $x^* = 0.6$ and $y^* = 0.4$. If the individual devotes a fraction 0.6 of its investment to the first behaviour and a fraction 0.4 of its investment to the second behaviour, then it obtains an average return of

$$a = 0.6\mu_1 + 0.4\mu_2 = 0.992.$$

The variance in return is obtained by noting that, when one splits allocation between two uncorrelated returns, R_1 and R_2, each with variance σ^2, the variance is

$$b = \mathrm{var}(xR_1 + yR_2) = [x^2 + y^2]\sigma^2.$$

Using the optimal split 0.6 vs. 0.4 for x and y, and the value $\sigma^2 = 0.1$ above, the variance is $b = 0.052$. The geometric mean is now approximately

$$a - b/2 = 0.992 - 0.052/2 = 0.966.$$

This mixture of behaviours therefore returns a higher geometric mean of 0.966 than when all investment is devoted to the higher yielding first behaviour, which has a geometric mean of 0.95, or when all investment is devoted to the lower yielding behaviour, which has a geometric mean of 0.93. This example illustrates the benefit of diversification when success multiplies over time.

Box 4: Reviews and technical issues

Gillespie (1994) and Ewens (2010) provide excellent technical overviews of genetic theory for variable environments. Recent reviews (Hedrick, 2006; Proulx & Adler, 2010) and new theory (Taylor, 2008) continue to appear.

Lande (2008) developed a comprehensive approach to the theory of fluctuating selection. He emphasized an adaptive topography method arising from the key insight that expected fitness depends on average reproductive success minus the covariance of reproductive success with population mean success. That expression for fitness is the same as developed in eqn (4), following from Gillespie (1977) and Frank & Slatkin (1990).

Lande (2008) also summarized the complexities that arise in diploid genetic systems under fluctuating selection. In general, diploid and multilocus models require one to pay attention to two issues (Frank & Slatkin, 1990). First, how do the fluctuations contributed by different alleles combine within individuals to determine the average and variance in individual reproductive success? Second, how do multiple alleles per individual induce correlations in reproductive success between copies of alleles in different individuals? The specific results that I gave in the text concern haploid models. Diploid and multilocus models must account for these additional complexities.

Rice (2008) introduced an exact expression for evolutionary change in stochastic models by expanding the scope and interpretation of the Price Equation. Tuljapurkar *et al.* (2009) review the theory of variable environments in relation to demography and life-history evolution. For economics theory, introductory microeconomics texts include overviews of the theory of risk and uncertainty. For investments, see Markowitz (1991). Okasha (2011) provides an entry to the philosophical literature on the theory of uncertainty in evolution and economics.

Note that reproductive returns are linear in this example. The entire benefit of diversification arises from the multiplicative nature of long-term success, which discounts variance.

The limitation of using individual geometric mean success

I used the individual's geometric mean success in this section. That assumption is valid only when we are interested in an absolute measure of an individual's long-term success in the absence of competition and relative comparison with others. However, it often does not make sense to measure success independently of others. Evolutionary success depends on the relative contribution to the population by a heritable type. The earlier sections of this article showed several different measures of success that arise from the temporal and spatial structure of competition and from the correlations in success between different types of individuals.

Suppose, for example, that the correlation between individual copies of an allele is low, $\rho \to 0$, as in the model of developmental variation in which $\rho = 1/N$ and population size, N, is large. Then, from eqn (9), we see that natural selection favours the type with the highest arithmetic average, μ, independently of the individual variance in reproductive success, σ^2. In that case, it does not make sense to analyse the geometric mean of individual reproductive success. Instead, success depends almost entirely on the arithmetic mean return taken over the two traits. If returns per trait are linear, $f'' = 0$, then the arithmetic mean is

$$x\mu_1 + y\mu_2 = x + y(1 - \gamma).$$

In this case, individuals will be favoured to allocate all resources to the higher yielding trait, labelled as trait one in this example.

Economic theories of variability and risk

Economic theories of risk and uncertainty typically focus on the absolute success of individuals or single agents. Relative success in economics concerns market share (Frank, 1990). However, there seems to be little economic theory about risk and uncertainty in relation to market share. Problems of market share lead to many issues discussed in this article. For example, relative success induces diminishing returns. The temporal and spatial scale of competition determines the proper measure of success.

One must also consider the proper unit of analysis to measure success and dynamics. If one is interested in the absolute currency value accumulated by an individual investor over a long period of time, then the individual's geometric mean success in return over successive intervals is often a good measure. If one is interested in an individual's purchasing power, then one must track the individual's currency valuation relative to the currency valuation among the population of individuals competing for the same goods.

If the individual has only a small fraction of the total pool of goods, then the individual's geometric mean return provides a good measure of success. However, if one is tracking a corporation or agent that controls a large fraction of the total resource pool, then the correlation between individual and total success may have a significant impact on outcome.

In some economic analyses, one is interested in behaviours or strategies. For example, what is the relative success of those following a particular financial strategy in the investment markets? The answer depends in part on whether all individuals following the same strategy have highly correlated returns or uncorrelated returns. A high correlation in returns between individuals following a strategy increases the variance in the aggregate success of that strategy. Higher variance usually leads to lower long-term success.

The frequency of the competing strategies also matters. Relatively rare strategies have a low correlation with the population average level of success, providing a rare-type advantage. The consequences of the rare-type advantage depend on whether competition occurs globally or locally over a series of isolated markets. These problems of relative success have received little attention in the study of economic competition.

Conclusions

Nearly all aspects of success include a variable component. To understand the consequences of that variability, one must study the hierarchical structure of traits within populations. Each individual has multiple traits. Each genotype or strategy has multiple individuals. Fluctuation in the success of particular traits has consequences that depend on the correlations between traits and the correlations between individuals. The aggregate variability of competing types affects relative success in ways that depend on density-dependent competition, which causes diminishing returns and induces an intrinsic frequency dependence that tends to favour rare types over common types.

The extensive biological theory of variability has dealt with particular aspects of the overall problem. But few analyses have set out the entire range of fluctuations, how those fluctuations are structured in populations, and the particular nature of competition that shapes the consequences of fluctuations. By considering the structure of the entire problem, one obtains a richer understanding of biological fitness and its consequences for evolutionary dynamics.

Many of the biological problems of variability also arise in economics. The theoretical literature in economics made the first analyses of success when there is a variable component of returns. But the biological literature has advanced further in the analysis of variability, particularly with respect to the importance of relative success and the hierarchical structure of competing types or strategies in populations.

Acknowledgments

I thank Sylvain Gandon for helpful comments. Parts of this article derive from Frank & Slatkin (1990). My research is supported by National Science Foundation grant EF-0822399, National Institute of General Medical Sciences MIDAS Program grant U01-GM-76499, and a grant from the James S. McDonnell Foundation.

References

Bernstein, W.J. & Wilkinson, D. 1997. Diversification, rebalancing, and the geometric mean frontier. Social Science Research Network. Available from http://ssrn.com/abstract=53503

Cooper, W.S. & Kaplan, R.H. 1982. Adaptive 'coin-flipping': a decision-theoretic examination of natural selection for random individual variation. *J. Theor. Biol.* **94**: 135–151.

Dempster, E.R. 1955. Maintenance of genetic heterogeneity. *Cold Spring Harbor Symp. Quant. Biol.* **20**: 25–32.

DeWitt, T.J. & Scheiner, S.M. 2004. *Phenotypic Plasticity: Functional and Conceptual Approaches*. Oxford University Press, New York.

Donaldson-Matasci, M.C., Lachmann, M. & Bergstrom, C.T. 2008. Phenotypic diversity as an adaptation to environmental uncertainty. *Evol. Ecol. Res.* **10**: 493–515.

Ewens, W.J. 2010. *Mathematical Population Genetics: I. Theoretical Introduction*, 2nd edn. Springer-Verlag, New York.

Frank, S.A. 1990. When to copy or avoid an opponent's strategy. *J. Theor. Biol.* **145**: 41–46.

Frank, S.A. & Slatkin, M. 1990. Evolution in a variable environment. *Am. Nat.* **136**: 244–260.

Gillespie, J.H. 1973. Polymorphism in random environments. *Theoret. Pop. Biol.* **4**: 193–195.

Gillespie, J.H. 1974a. Natural selection for within-generation variance in offspring number. *Genetics* **76**: 601–606.

Gillespie, J.H. 1974b. Polymorphism in patchy environments. *Am. Nat.* **108**: 145–151.

Gillespie, J.H. 1977. Natural selection for variances in offspring numbers: a new evolutionary principle. *Am. Nat.* **111**: 1010–1014.

Gillespie, J.H. 1978. A general model to account for enzyme variation in natural populations. V. The SAS-CFF model. *Theoret. Pop. Biol.* **14**: 1–45.

Gillespie, J.H. 1994. *The Causes of Molecular Evolution*. Oxford University Press, New York, NY, USA.

Grafen, A. 1999. Formal darwinism, the individual-as-maximizing-agent analogy and bet-hedging. *Proc. R. Soc. Lond. B* **266**: 799–803.

Hartl, D.L. & Cook, R.D. 1973. Balanced polymorphisms of quasineutral alleles. *Theoret. Pop. Biol.* **4**: 163–172.

Hedrick, P.W. 2006. Genetic polymorphism in heterogeneous environments: the age of genomics. *Annu. Rev. Ecol. Evol. Syst.* **37**: 67–93.

Houston, A.I. & McNamara, J.M. 1992. Phenotypic plasticity as a state-dependent life-history decision. *Evol. Ecol.* **6**: 243–253.

Karlin, S. & Lieberman, U. 1974. Random temporal variation in selection intensities: case of large population size. *Theoret. Pop. Biol.* **6**: 355–382.

King, O.D. & Masel, J. 2007. The evolution of bet-hedging adaptations to rare scenarios. *Theoret. Pop. Biol.* **72**: 560–575.

Kussell, E. & Leibler, S. 2005. Phenotypic diversity, population growth, and information in fluctuating environments. *Science* **309**: 2075–2078.

Lande, R. 2008. Adaptive topography of fluctuating selection in a mendelian population. *J. Evol. Biol.* **21**: 1096–1105.

Levene, H. 1953. Genetic equilibrium when more than one ecological niche is available. *Am. Nat.* **87**: 331–333.

MacBeth, J. D. 1995. What's the long-term expected return to your portfolio? *Financial Analysts Journal* **51**: 6–8.

Markowitz, H. M. 1991. *Portfolio Selection: Efficient Diversification of Investments*, 2nd edn. Blackwell, Malden, MA, USA.

McNamara, J.M. 1995. Implicit frequency dependence and kin selection in fluctuating environments. *Evol. Ecol.* **9**: 185–203.

Moran, N.A. 1992. The evolutionary maintenance of alternative phenotypes. *Am. Nat.* **139**: 971–989.

Okasha, S. 2011. Optimal choice in the face of risk: decision theory meets evolution. *Philos. Sci.* **78**: 83–104.

Orr, H.A. 2007. Absolute fitness, relative fitness, and utility. *Evolution* **61**: 2997–3000.

Pigliucci, M. 2001. *Phenotypic Plasticity: Beyond Nature and Nurture*. Johns Hopkins University Press, Baltimore, MD, USA.

Proulx, S.R. & Adler, F.R. 2010. The standard of neutrality: still flapping in the breeze? *J. Evol. Biol.* **23**: 1339–1350.

Real, L. 1980. Fitness, uncertainty, and the role of diversification in evolution and behavior. *Am. Nat.* **115**: 623–638.

Rice, S.H. 2008. A stochastic version of the Price equation reveals the interplay of deterministic and stochastic processes in evolution. *BMC Evol. Biol.* **8**: 262.

Sasaki, A. & Ellner, S. 1995. The evolutionarily stable phenotype distribution in a random environment. *Evolution* **49**: 337–350.

Scheiner, S.M. 1993. Genetics and evolution of phenotypic plasticity. *Annu. Rev. Ecol. Syst.* **24**: 35–68.

Seger, J. & Brockmann, H.J. 1987. What is bet-hedging? *Oxford Surv. Evol. Biol.* **4**: 182–211.

Stearns, S.C. 1989. The evolutionary significance of phenotypic plasticity. *Bioscience* **39**: 436–445.

Stephens, D. W. & Krebs, J. R. 1986. *Foraging Theory*. Princeton University Press, Princeton.

Taylor, J.E. 2008. Environmental variation, fluctuating selection and genetic drift in subdivided populations. *Theoret. Pop. Biol.* **74**: 233–250.

Tuljapurkar, S., Gaillard, J.M. & Coulson, T. 2009. From stochastic environments to life histories and back. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **364**: 1499.

Via, S., Gomulkiewicz, R., De Jong, G., Scheiner, S., Schlichting, C. & Van Tienderen, P. 1995. Adaptive phenotypic plasticity: consensus and controversy. *Trends Ecol. Evol.* **10**: 212–217.

West-Eberhard, M.J. 2003. *Developmental Plasticity and Evolution*. Oxford University Press, New York.

Wilbur, H.M. & Rudolf, V.H.W. 2006. Life-history evolution in uncertain environments: bet hedging in time. *Am. Nat.* **168**: 398–411.

Received 20 May 2011; revised 30 July 2011; accepted 4 August 2011

JOURNAL OF Evolutionary Biology

doi: 10.1111/j.1420-9101.2011.02373.x

REVIEW

Natural selection. II. Developmental variability and evolutionary rate*

S. A. FRANK

Department of Ecology and Evolutionary Biology, University of California, Irvine, CA, USA

Keywords:

adaptive dynamics;
Baldwin effect;
genetic assimilation;
natural selection;
phenotypic plasticity;
population genetics;
rapid adaptation;
theory.

Abstract

In classical evolutionary theory, genetic variation provides the source of heritable phenotypic variation on which natural selection acts. Against this classical view, several theories have emphasized that developmental variability and learning enhance nonheritable phenotypic variation, which in turn can accelerate evolutionary response. In this paper, I show how developmental variability alters evolutionary dynamics by smoothing the landscape that relates genotype to fitness. In a fitness landscape with multiple peaks and valleys, developmental variability can smooth the landscape to provide a directly increasing path of fitness to the highest peak. Developmental variability also allows initial survival of a genotype in response to novel or extreme environmental challenge, providing an opportunity for subsequent adaptation. This initial survival advantage arises from the way in which developmental variability smooths and broadens the fitness landscape. Ultimately, the synergism between developmental processes and genetic variation sets evolutionary rate.

> In evolutionary biology, environmentally induced modifications come under unfinished business… There have been repeated assertions of both their importance and their triviality, a lot of discussion with no consensus… Yet the debate has continued over such concepts as genetic assimilation, the Baldwin effect, organic selection, morphoses, and somatic modifications. So much controversy over the span of a century suggests that a problem of major significance remains unsolved (West-Eberhard, 2003, p. 498).

Introduction

A single genotype produces different phenotypes. Developmental programs match the phenotype to different environments. Intrinsic developmental fluctuations spread the distribution of phenotypes. Extrinsic environmental fluctuations perturb developmental trajectory. These nonheritable types of phenotypic variation are common.

Nonheritable phenotypic variation is not transmitted through time. Thus, nonheritable variation would seem to be irrelevant for evolutionary change, which instead depends on the genetic component of variation. However, nonheritable phenotypic variation can, in principle, affect evolutionary rate. At first glance, that contribution of nonheritable phenotypic variation to evolutionary rate appears to be a paradox.

Many different theories, commentaries, and controversies turn on this paradox (Box 2). The literature has followed a consistent pattern. Detailed theories relate developmental variability to accelerated evolution. Counterarguments ensue. Listings of complicated examples claim to support the theory. Refinements to the theory develop.

In the end, few compelling examples relate nonheritable phenotypic variability to evolutionary rate. The literature is hard to read. Enthusiasts extend the concepts and keep the problem alive. Through the enthusiasts' promotions, many have heard of the theory. But, in practice, few consider the role of nonheritable phenotypic variability in their own analyses of evolutionary rate. Almost everyone ignores the problem.

In this article, I emphasize simple theory that relates nonheritable phenotypic variability to evolutionary rate. Understanding the paradoxical relation between nonheritable phenotypic variability and evolutionary rate is an essential step in reasoning about many evolutionary problems.

Correspondence: Steven A. Frank, Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697–2525, USA.
Tel.: +1 949 824 2244; fax: +1 949 824 2181; e-mail: safrank@uci.edu
*Part of the Topics in Natural Selection series (see Box 1).

Box 1: Topics in the theory of natural selection

This article is part of a series on natural selection. Although the theory of natural selection is simple, it remains endlessly contentious and difficult to apply. My goal is to make more accessible the concepts that are so important, yet either mostly unknown or widely misunderstood. I write in a nontechnical style, showing the key equations and results rather than providing full derivations or discussions of mathematical problems. Boxes list technical issues and brief summaries of the literature.

This article is primarily a concise tutorial to the basic concepts (see Box 1). I briefly mention some of the history (Box 2) and recent, more advanced literature (Box 3).

Smoothing the evolutionary path

The distribution of phenotypes for a given genotype is called the *reaction norm*. All theories come down to the fact that a broad reaction norm smooths the path of increasing fitness. Once one grasps the smoothing process, many apparently different theories become easy to understand.

The next section gives the mathematical expression for the smoothing of fitness by the reaction norm. Figure 1 explains the mathematics with a simple example.

The reaction norm smooths fitness

We need to track three quantities. First, fitness, $f(x)$, varies according to the particular phenotype expressed, x.

Second, the phenotype expressed varies according to the reaction norm. Read $p(x|\bar{x})$ as the probability of expressing the phenotype x given a genotype with average phenotype $\bar{x}$.

Third, we must calculate $F(\bar{x})$, the expected fitness for a genotype with average phenotype $\bar{x}$. We obtain the expected fitness by summing up the probability, p, of expressing each phenotype multiplied by the fitness, f, of each phenotype. That sum is

$$F(\bar{x}) = \sum p(x|\bar{x})f(x), \tag{1}$$

taken over all the different phenotypes, x. We often measure x as a continuous variable. The sum is then equivalently written as

$$F(\bar{x}) = \int p(x|\bar{x})f(x)dx. \tag{2}$$

This equation shows how one averages the fitness, $f(x)$, for each phenotypic value, x, over the reaction norm, $p(x|\bar{x})$, to obtain the expected fitness of a genotype, $F(\bar{x})$. We label each genotype by its average phenotype, $\bar{x}$. The expected fitness of a genotype, $F(\bar{x})$, is what matters for evolutionary process (see Frank, 2011, for the role of variability in fitness).

The averaging of expected fitness over the reaction norm is the key to the entire subject. Averaging over the reaction norm, p, flattens and smooths the fitness function, f. This smoothing makes the curve for expected fitness, F, have lower peaks and shallower valleys than the original fitness curve, f. The smoothing of F changes evolutionary dynamics. The whole problem comes down to understanding how reaction norms smooth fitness, and the consequences of a smoother relation between genotype and fitness.

Box 2: Historical overview

Schlichting & Pigliucci (1998) and West-Eberhard (2003) thoroughly review the subject. Here, I highlight a few key points in relation to this article. I treat learning and developmental plasticity as roughly the same with regard to potential consequences for evolutionary rate, although one could certainly choose to focus on meaningful distinctions.

In my own reading during the 1980s, I had found the relation between learning and evolutionary rate intriguing but confusing. Baldwin's (1896) idea that learning can accelerate evolutionary rate seemed attractive. Mayr (1982), in his monumental review of biological thought, also discussed various ways in which behavior or flexible developmental programs might alter evolutionary dynamics. Those ideas seemed potentially important, but it was not easy to grasp the essence. The literature at that time was not helpful, with a lot of jargon and sometimes almost mystical commentary mixed in with intriguing and creative ideas.

It was clear that learning could slow evolutionary rate. Different genotypes could, through learning, end with the same phenotype. Reducing the phenotypic distinction between different genotypes would generally slow evolutionary rate. The more intriguing problem concerns the origin of evolutionary novelty or the response to novel or extreme environmental challenge. Environmental novelty and acceleration of evolutionary response were the primary concern of Baldwin (1896), Waddington (1942, 1953), and West-Eberhard (2003). My article also focuses on acceleration of evolutionary response.

Hinton & Nowlan (1987) clarified the subject with their simple conclusion that:

> Learning alters the shape of the search space in which evolution operates and thereby provides good evolutionary paths towards sets of co-adapted alleles. We demonstrate that this effect allows learning organisms to evolve *much* faster than their nonlearning equivalents, even though the characteristics acquired by the phenotype are not communicated to the genotype.

During the past few decades, the fundamental role of smoothed fitness surfaces in biology has not always been recognized as fully as it should be, in spite of several fine papers along that line (see Box 3). Interestingly, certain computer optimization algorithms take advantage of the increased search speed provided by a process similar to smoothed fitness landscapes (Kirkpatrick *et al.*, 1983; Geyer & Thompson, 1995).

Box 3: Recent literature

Ancel (2000) analyzed smoothed fitness surfaces and the consequences for evolutionary rate. She emphasized three important points.

First, learning accelerates evolution only under certain conditions. The examples in the text illustrate this point by showing that learning mainly accelerates evolution through discovery of viable phenotypes or in the smoothing of a multipeaked fitness surface. Otherwise, the smoothing of fitness surfaces may lower the maximum fitness that can be attained, reducing the slope and the evolutionary rate to the peak.

Second, although learning may accelerate evolution, it is not necessarily true that learning evolved because it accelerates evolution. The evolutionary consequence of a trait is distinct from whether or not the trait evolved because of its potential to alter subsequent evolutionary dynamics. The literature discusses this distinction under the topic of evolvability. Evolvability has developed into a large subject of its own (e.g. Wagner & Altenberg, 1996; Kirschner & Gerhart, 1998; Pigliucci, 2008; Rajon & Masel, 2011; Woods *et al.*, 2011). Holland's (1975) distinction between exploration and exploitation captures aspects of the later developments on evolvability.

Third, Ancel noted historical precedents for the idea that phenotypic variance may eliminate otherwise uncrossable valleys in fitness landscapes (Wright, 1931; Lande, 1980; Whitlock, 1997).

A large literature develops issues related to Ancel's three points and the broader problems of how reaction norms affect fitness surfaces. I list a small sample (De Jong, 1990; Gavrilets & Scheiner, 1993; Anderson, 1995; Frank, 1996; Turney, 1996; Mayley, 1997; Turney *et al.*, 1997; Pigliucci, 2001; Rice, 2002; Hall, 2003; Price *et al.*, 2003; Gavrilets, 2004; Mills & Watson, 2006; Pigliucci *et al.*, 2006; Crispo & Rausher, 2007; Suzuki & Arita, 2007; Lande, 2009; Chevin & Lande, 2010; Chevin *et al.*, 2010; Gavrilets, 2010).

West-Eberhard (2003) discusses many empirical issues and examples. Recent studies in evolutionary biology provide new data or summaries of the literature (Aubret & Shine, 2009; Bell & Robinson, 2011). Articles in microbiology and cancer research have also developed the relation between nonheritable phenotypic variation and evolutionary process (Rubin, 1990; Booth, 2002; Sumner & Avery, 2002; Yomo *et al.*, 2005; Avery, 2006; Niepel *et al.*, 2009; Spencer *et al.*, 2009; Altschuler & Wu, 2010; Kaneko, 2011).

In the text, I discuss the synergism between genetic and developmental variation. I am not aware of literature related to that issue. However, given the many papers on the general topic, the synergism between genetics and development may have come up previously.

Example of continuous smoothing

Figure 1 shows an example of smoothing with discrete distributions. It will often be convenient to consider smoothing of continuous variables. Figure 2 shows an example. The following expressions describe the underlying mathematics.

In Fig. 2, the reaction norm follows a normal distribution. In symbols, we write

$$p(x|\bar{x}) \sim \mathcal{N}(\bar{x}, \gamma^2),$$

which we read as the probability, p, of a phenotype, x, for a reaction norm centered at $\bar{x}$, follows a normal distribution with mean $\bar{x}$ and variance γ^2.

For fitness, we write in symbols

$$f(x) \sim \mathcal{N}(0, \sigma^2),$$

which we read as the fitness, f, of a phenotype, x, has the shape of a normal distribution with mean 0 and variance σ^2. In this case, we assume the center of the fitness distribution is at a phenotypic value of zero to give a fixed point for comparison – any value to center fitness could be used. The important issue is that fitness falls off from its peak by the pattern of a normal distribution. The width of the fitness function is set by the variance parameter, σ^2.

We can now use eqn 2 to calculate the expected fitness of a genotype with average phenotype $\bar{x}$, yielding

$$F(\bar{x}) \sim \mathcal{N}(0, \gamma^2 + \sigma^2). \quad (3)$$

This equation shows that smoothing by the reaction norm, p, flattens and widens the shape of the fitness function by increasing the variance of the expression for F.

Evolutionary response to novel or extreme challenge

If a genotype expresses an average phenotype close to the maximum fitness, then a narrow reaction norm has higher fitness than a broad reaction norm. The lower plots of $F(\bar{x})$ in Fig. 2 illustrate contrasting widths of reaction norms. Near the peak, the average phenotype closely matches the optimum, and the narrower reaction norm has higher fitness. This advantage occurs because a narrow reaction norm expresses fewer phenotypes in the tails, away from the optimum.

For genotypes with an average phenotype far from the maximum fitness, a broad reaction norm has higher fitness than a narrow reaction norm. Figure 3 illustrates this advantage for broad reaction norms. In that figure, both reaction norms are centered at $\bar{x}$. Only those phenotypes above the fitness truncation point survive. The broad reaction norm produces some individuals with phenotypes above the truncation point, whereas the narrow reaction norm has zero fitness.

If the environment poses a novel or extreme challenge, the broad reaction norm wins. By contrast, in a stable environment for which the current average phenotype is close to the fitness optimum, the narrow reaction norm wins. Thus, extreme or novel environmental challenges or intense competition favor a broad reaction norm.

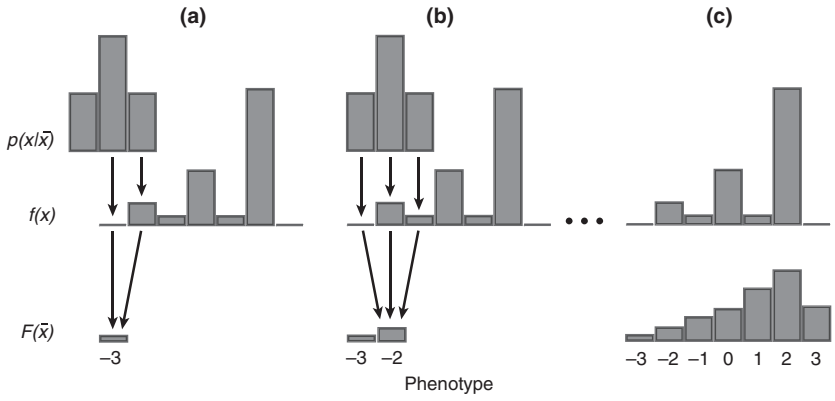

Fig. 1 The reaction norm smooths the fitness landscape. This simple example illustrates the calculation of the expected fitness for each genotype, following eqn 1. (a) The calculation of expected fitness, $F(\bar{x})$, for the smallest average phenotype, $\bar{x} = -3$. For that average phenotype, the reaction norm, $p(x|\bar{x})$, shows the probabilities of expressing different phenotypes, x. In this case, the peak of the reaction norm matches the average value, and each phenotype ± 1 occurs half as often as the peak value. To get the expected fitness for a reaction norm centered at $\bar{x} = -3$, one sums up the probability $p(x|\bar{x})$ for each phenotype, x, multiplied by the fitness for each phenotype, $f(x)$. The arrows illustrate the summation. (b) The expected fitness, $F(\bar{x})$, for each increase in $\bar{x}$, is calculated by the same summation process, shifting the reaction norm to the right by one to get the proper value for each $\bar{x}$. (c) The full transformation is shown between the fitness for each phenotypic value, $f(x)$, and the expected fitness, $F(\bar{x})$, for each genotype with reaction norm $p(x|\bar{x})$ and average phenotype $\bar{x}$. The reaction norm smooths the multipeaked fitness function, $f(x)$, into the single-peaked fitness function $F(\bar{x})$. Evolutionary dynamics depend on genotypic fitnesses, F. Thus, the reaction norm transforms fitness into a smooth function that allows a direct increasing path to the fitness peak from any starting value for average phenotype.

Haldane (1932) made a similar point when he said: 'Intense competition favors variable response to the environment rather than high average response. Were this not so, I expect that the world would be much duller than is actually the case'. Holland's (1975) emphasis on exploration versus exploitation is perhaps closer to the problem here. Broad reaction norms are favored when exploration of novel challenges dominates, whereas narrow reaction norms are favored when exploitation dominates. Fluctuating environments may also favor a broad reaction norm to increase the chance of matching whatever is favored at any time (Frank, 2011). Here, I focus on constant challenges to extreme or novel environments.

Smoothly increasing fitness path in a multipeak fitness landscape

Much discussion in evolutionary theory concerns how populations shift from a lower fitness peak to a higher fitness peak (Coyne *et al.*, 1997). For example, in the fitness landscape, $f(x)$, of Fig. 4b, a population starting on a lower peak must evolve through a valley of lower fitness in order to follow an increasing path to a higher fitness peak. Natural selection typically follows a path of increasing fitness, so a population may be trapped on a lower peak.

Most evolutionary analyses use a fitness landscape that relates phenotype, x, to fitness, $f(x)$. However, the proper measure should relate the average phenotype of a genotype, $\bar{x}$, to the expected fitness, $F(\bar{x})$ (here ignoring variation in fitness, Frank, 2011).

A sufficiently broad reaction norm smooths a multipeak fitness landscape, $f(x)$, into a smooth landscape, $F(\bar{x})$, with a single peak (Fig. 4c). A broad reaction norm will typically perform badly near a fitness peak, but allow much more rapid evolutionary advance to a higher fitness peak. Once again, we see that broad reaction norms exploit current fitness opportunities relatively poorly but gain by enhanced exploration and achievement of novel adaptations.

Dimensionality and discovery

The reaction norm may be generated randomly by perturbations in development. If so, then exploration of the fitness landscape by a broad reaction norm is a type of random search. Figures 3 and 4 show that random search can greatly increase the rate of adaptation, particularly to novel environmental challenges.

Those previous examples showed the reaction norm and fitness both varying across a single dimension. A broad reaction norm spreads phenotypes along that single dimension, increasing the chance that some individuals will have high fitness.

Now consider the much more difficult search problem that arises in higher dimensions (Gavrilets, 2004). Suppose, for example, that adapting to a novel environmental challenge requires multiple phenotypic changes to work together in a harmonious way. Think of each

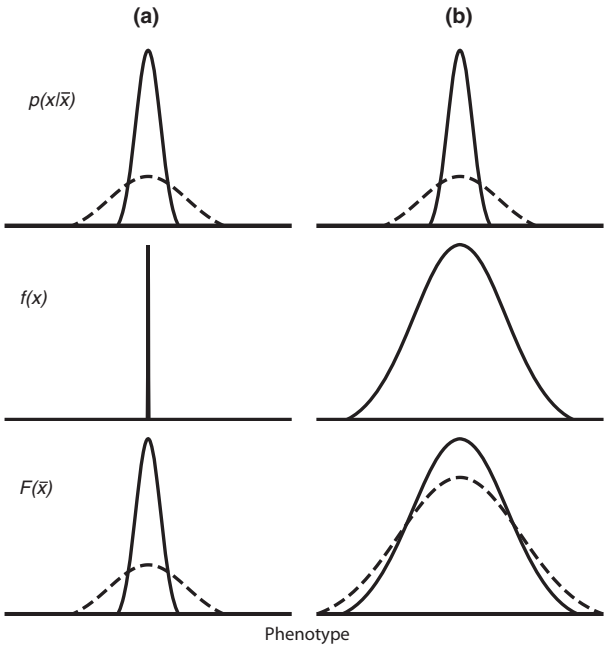

Fig. 2 Reaction norms and fitness for continuous phenotypes. Each column shows how the reaction norm, $p(x|\bar{x})$, smooths the fitness function, $f(x)$, to give the expected fitness, $F(\bar{x})$, for a genotype with average phenotype $\bar{x}$. The smoothing follows eqn 2. These examples use normal distributions that lead to eqn 3. (a) The solid and dashed reaction norms follow $\mathcal{N}(\bar{x}, 1/2)$ and $\mathcal{N}(\bar{x}, 5)$, respectively. Fitness, $f(x)$, has the shape of a normal distribution with vanishingly small variance, $\mathcal{N}(0, \sigma^2 \to 0)$. Thus, expected fitness, $F(\bar{x})$, is the same as the reaction norm. (b) The same structure as in (a), except that $f(x)$ is much wider, following $\mathcal{N}(0, 7)$. Thus, $F(\bar{x})$ now has curves $\mathcal{N}(0, 7.5)$ and $\mathcal{N}(0, 12)$ for solid and dashed curves, respectively. In each plot, the baseline is set to 4.3% of the peak in that plot. The baseline truncates phenotypes with low vigor, setting their fitnesses to zero.

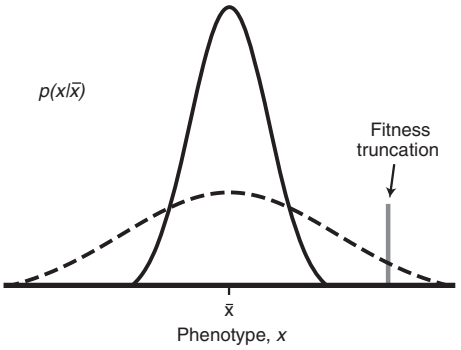

Fig. 3 Novel environmental challenge or intense competition favors a broad reaction norm. In this example, both the broad and narrow reaction norms are centered at $\bar{x}$. Phenotypes above the truncation point survive. Phenotypes below the truncation point die. None of the phenotypes for the narrow reaction norm are above the truncation point, so all die. Some of the phenotypes of the broad reaction norm survive. Those surviving phenotypes may evolve so that their average phenotype, $\bar{x}$, moves toward the truncation point, improving fitness over time. Improvement occurs if there is genetic variation for the average phenotype, $\bar{x}$, of the reaction norm.

particular phenotypic change as a trait in its own dimension, so that the search now occurs in multiple dimensions. If the reaction norm simply generates random phenotypes in each dimension, then there is little chance of getting simultaneous matching phenotypes in multiple dimensions.

To visualize the multidimensional problem, begin with the one-dimensional fitness landscape in Fig. 4b. Now consider two phenotypic dimensions. Assume that fitness concentrates along one dimension, as in Fig. 5a. In that plot, only a narrow band of phenotypes along the second phenotypic dimension produces viable individuals. In the first dimension, fitness rises and falls along the same peaks and valleys as in Fig. 4b. Thus, both figures show essentially the same fitness landscape, but in the second case the nearly one-dimensional landscape is embedded in a second dimension (fitnesses scale logarithmically in Fig. 5).

In two dimensions, the reaction norm will smooth phenotypes along both trait axes. When the reaction norm varies mostly along the same dimension as the variation in fitness, as in Fig. 5b, then we obtain the same smoothing as in one dimension (dashed curve of Fig. 4c). When the reaction norm varies in both directions, as in Fig. 5d, then the smoothed surface has very low fitness even at its peak. The low fitness occurs because the randomly generated reaction norm produces phenotypes spread across two dimensions. Most of those phenotypes fall off of the one dimensional concentration of fit phenotypes.

In general, when the dimensionality of the reaction norm exceeds the dimensionality of the fitness concentration, then a random search process is inefficient. The cost of exploration is so high that even the best average phenotype for a genotype has fitness, $F(\bar{x})$, lower than the lowest peak of the fitness landscape, $f(x)$. Here, $\bar{x}$ and x represent multidimensional phenotypes. Figure 6 illustrates the cost of exploration in relation to the spread across dimensions.

In summary, if the space of possible trait combinations spreads over greater dimensions than the concentration of fitness, then randomly generated variations will produce mostly worthless variants. The search cost is high, and average performance for a widely spread reaction norm is low. The smoothed fitness surface may have a steadily rising path to its fitness maximum from many initial points, but the height of the fitness peak is so low that a broad reaction norm will often be strongly selected against.

The following sections describe two processes that may offset the high cost of developmental variation. First, the broad search space may be covered by genetic variants, with developmental variation searching only the local regions around each genotypic variant.

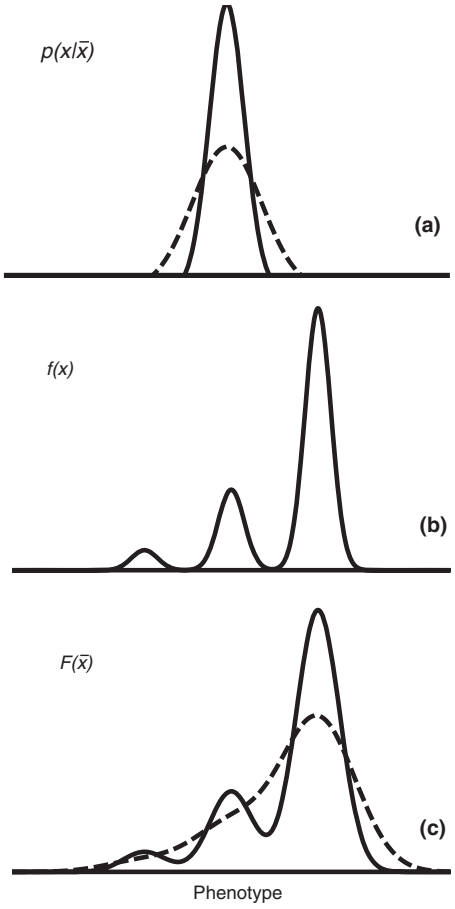

Fig. 4 A broad reaction norm smooths a multipeak fitness landscape. (a) The dashed curve shows the broader reaction norm, $p(x|\bar{x})$. (b) The fitness landscape for each particular phenotype, $f(x)$, has multiple peaks. (c) The broad reaction norm smooths the fitness landscape to a single peak for the relation between the average phenotype for a genotype, $\bar{x}$, and fitness, $F(\bar{x})$. In this example, the narrow and broad reaction norms follow $\mathcal{N}(0, \gamma^2)$ distributions with variances of 0.04 and 0.16, respectively. Fitness is given by $f(x) = \sum_{i=-1}^{1}(3|1 + i|^2 + 1)\mathcal{N}(i, \sigma^2)$, with $\sigma^2 = 0.0225$. The value of $F(\bar{x})$ is calculated from eqn 2, yielding the expression for $f(x)$ in the prior sentence with the variance replaced by $\sigma^2 + \gamma^2$. The baseline truncates small values.

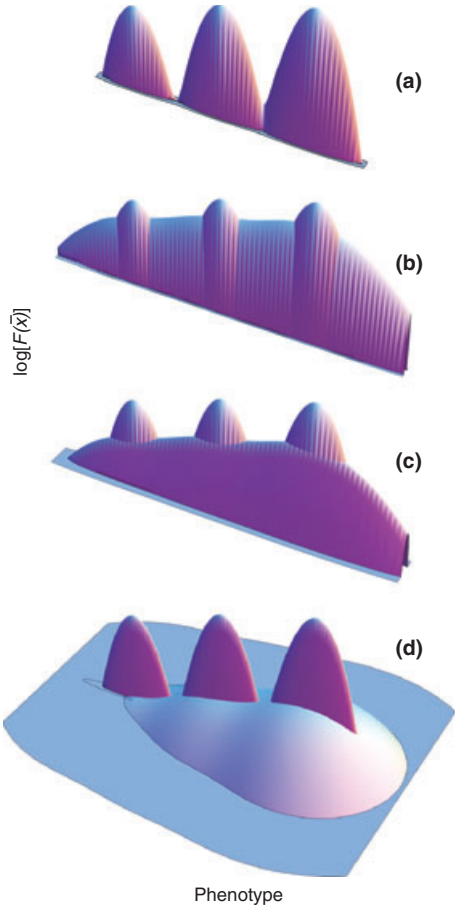

Fig. 5 A broad reaction norm performs poorly when fitness is concentrated in a lower dimension. (a) The bivariate analogy of the fitness landscape in Fig. 4b, scaled logarithmically. The primary dimension has variance $\sigma_1^2 = 0.0225$ corresponding to standard deviation $\sigma_1 = 0.15$, as in Fig. 4b. The secondary (narrow) dimension has standard deviation $\sigma_2 = 0.1\sigma_1$. (b) Fitness landscape smoothed by a reaction norm concentrated in the same dimension as fitness. The variance of the reaction norm in the primary dimension is $\gamma_1^2 = 0.16$, and standard deviation is $\gamma_1 = 0.4$, as in the dashed reaction norm of Fig. 4a. The standard deviation in the secondary dimension is $\gamma_2 = 0.01\gamma_1$. The smoothed fitness surface rises steadily to a peak along its ridge in the primary dimension, tracing the same path as the dashed curve in Fig. 4c. (c and d) Increasingly broad reaction norms in the secondary dimension with standard deviations of $0.1\gamma_1$ and γ_1, respectively. The baseline truncates small fitness values, which are considered inviable.

Second, developmental variation may be biased in a way that tends to match the environment. If a developmental or learning process brings the phenotype close to the concentration of fitness in a high dimensional space, then some additional random variation can greatly increase the rate of adaptation. In this case, the fitness surface is smoothed to provide a steady path of increasing fitness, and the developmental bias that brings the center of the phenotypic distribution close to the fitness concentration mitigates the large cost of search in high dimensional phenotypic spaces.

Synergism between phenotypic and genetic variation

A broad reaction norm may enhance survival and subsequent opportunity for improved fitness. But those benefits arise only when a genotype is sufficiently close to a fitness peak. Figure 7 illustrates the problem.

Figure 7a shows the fitness peak in a novel environment. The dots show the locations of alternative geno-

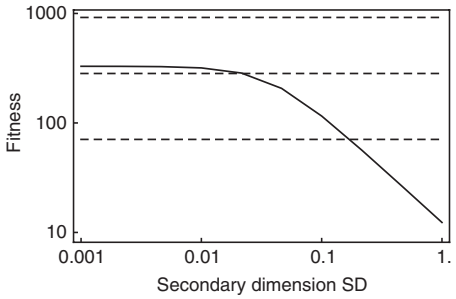

Fig. 6 Decline in fitness with an increasingly broad reaction norm away from the primary concentration of fitness. The dashed lines show the fitness associated with the high, medium and low peaks of the fitness landscape in Fig. 5a. The solid curve shows the highest point of the fitness functions smoothed by the reaction norms of Fig. 5b–d, with standard deviation of the reaction norm increasing in the secondary dimension. The secondary dimension standard deviation value shown in the plot gives the amount by which the primary dimension standard deviation is multiplied in the second dimension. The multipliers 0.01, 0.1, 1.0 correspond to the three smoothed fitness surfaces in Fig. 5b–d. When the secondary dimension is narrow, for example, reduced in width by a factor 0.01, then the smoothed fitness peak is higher than the intermediate fitness peak of the unsmoothed landscape, as in Fig. 4c. As width in the secondary dimension increases, the cost of exploring in a dimension away from the concentration of fitness causes the peak of the smoothed fitness landscape to drop very low, illustrating the very high cost of exploring in more dimensions than the concentration of fitness.

types, placed by their average phenotypes in two dimensions. Neither genotype has positive fitness. Both will die out. In that plot, the fitness peak is the direct fitness landscape, unsmoothed by a reaction norm. In Fig. 7b, the reaction norm is relatively narrow, smoothing the fitness landscape. But that smoothing is not enough to place either genotype on the nonzero fitness surface. Both genotypes still die out.

The broader reaction norm in Fig. 7c smooths the fitness surface more widely. That additional smoothing allows the nearby genotype to survive. Subsequent small genetic variations would allow natural selection to drive the surviving population up the path of increasing fitness to the fitness peak.

The distant (red) genotype cannot survive even with the broad reaction norm of Fig. 7c. The contrast between the nearby and distant genotypes emphasizes a key point. A genotype must be sufficiently close to the nonzero part of the smoothed fitness surface in order for the developmental variation of the reaction norm to allow survival – the touching of the fitness surface. If a genotype touches the fitness surface, then it can seed a population in which small genetic variations allow subsequent adaptation by climbing the surface to the peak.

In a high dimensional space, any single genotype is unlikely to be located sufficiently close to a fitness peak after a significant change in the environment or in response to an unpredictable challenge. Synergism between genetic variation and the phenotypic variation of reaction norms provides one solution to this search problem.

Figure 8 illustrates synergism between genetic and phenotypic variation. The dots represent different genotypes. Each genotype has a different combination of average phenotypic values in two dimensions. The array of dots shows the genetic diversity in the population. The

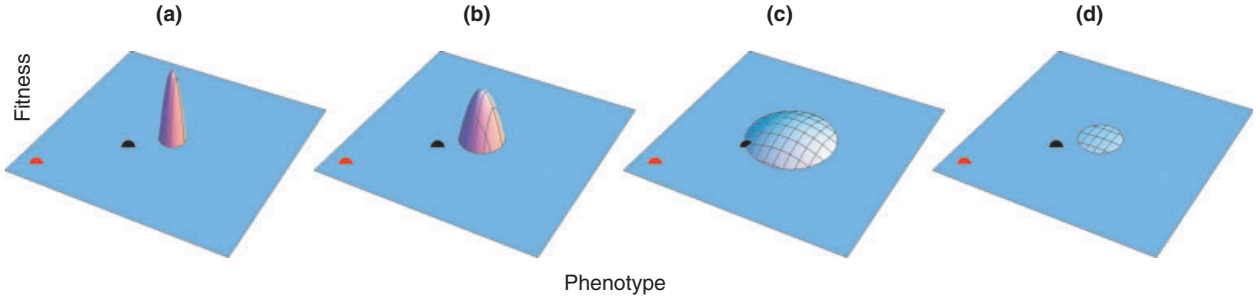

Fig. 7 Fitness in two phenotypic dimensions after challenge by a novel or extreme environment. The peak corresponds to the favored phenotype after environmental challenge. The black dot shows the average phenotype for a genotype near the new fitness peak. The red dot shows the average phenotype for a genotype relatively far from the new peak. A sufficiently broad reaction norm allows the nearby genotype to survive, providing an opportunity for natural selection to drive the population up the smoothly increasing path to the new fitness peak. By contrast, the distant genotype cannot survive the environmental challenge. (a) The fitness landscape, $f(x)$, showing the direct relation between phenotype and fitness when not smoothed by a reaction norm. (b) A relatively narrow reaction norm smooths the fitness peak, $F(\bar{x})$, but not sufficiently to allow the nearby genotype to survive. (c) A broader reaction norm allows the nearby genotype to survive, with subsequent opportunity for natural selection to drive the population to the peak. (d) An increasingly broad reaction norm causes the smoothed fitness peak to sink mostly below the fitness truncation level, so that the nearby genotype cannot survive. All plots show a bivariate normal fitness surface with mean (1/2,1/2) and variance $\sigma^2 + \gamma^2$, with $\sigma^2 = 0.01$ for the fitness landscape, and γ^2 for the reaction norm of 0, 0.01, 0.07, 0.11 for plots left to right. The heights are the natural logarithm of fitness, with a truncation base of log (10). The nearby black dot is at (17/40,17/40), and the far red dot is at (11/40,11/40).

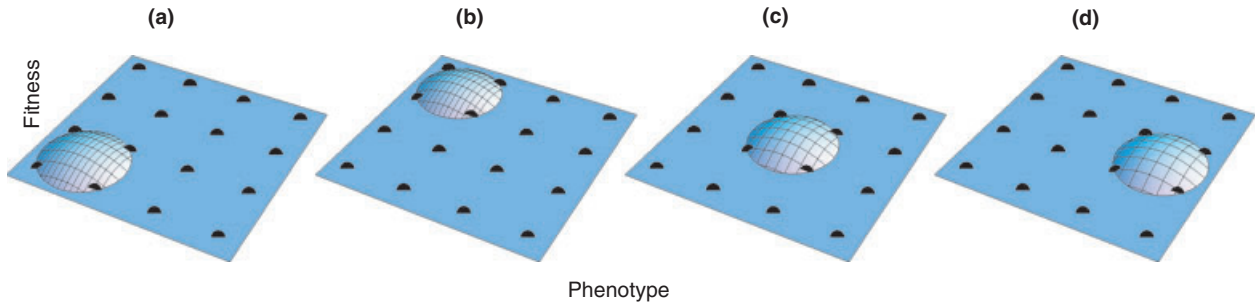

Fig. 8 Synergism between the reaction norm and genetic variation allows rapid adaptation to novel or extreme environments. In this case, a population has multiple genotypes, each genotype located at one of the dots. The smoothed fitness surface is sufficiently broad that, for any location of the fitness peak after environmental challenge, the fitness surface touches at least one of the genotypes. The genotypes that touch the fitness surface survive, allowing the surviving population the potential subsequently to evolve up the fitness surface to the new peak. The dots are located at all bivariate pairs from $\{(5 + 6i)/40,(5 + 6j)/40\}$ for $i,j = 1,\ldots,4$. The fitness surface has variance $\sigma^2 + \gamma^2 = 0.08$, as in Fig. 7c.

smoothed fitness surface has the same fitness peak and reaction norm shape as in Fig. 7c. In Fig. 8, the location of the fitness surface varies in the different plots, illustrating different environmental challenges. No matter where the newly favored fitness surface arises upon environmental challenge, the genetic diversity in the population provides at least one genotype on the nonzero part of the novel fitness surface. Those genotypes on the surface can survive the novel challenge. Subsequent small genetic variations around a surviving genotype allow the population to evolve up the fitness surface to the peak set by the novel environmental challenge.

Synergism between genetic and phenotypic variation divides the adaptive search problem into three parts. Genetic variation covers widely separated locations in the phenotype space. Reaction norms cover the phenotype space around each genotype. Any genotype on a nonzero part of a novel fitness surface can survive and subsequently adapt by small genetic variations and natural selection. See Box 4 for an example of the synergism between genetic and nonheritable phenotypic variation.

Matching the environment by plasticity or learning

To survive a novel environmental challenge, a phenotype must be near the nonzero part of the new fitness landscape. A population may survive by having a variety of genotypes that produce different phenotypes, increasing the chance that at least one of the phenotypes will be close to a new fitness peak. Alternatively, a single genotype may be able to produce diverse phenotypes by matching phenotypic expression to the particular environment. The developmental flexibility to match environments may arise by phenotypic plasticity or learning.

Plasticity or learning may not be able to match exactly a novel or extreme environmental challenge. But if a developmental response to the environment can move the phenotype sufficiently close to the nonzero part of the new fitness landscape, then the genotype may survive and subsequently adapt (Baldwin, 1896; Waddington, 1942; West-Eberhard, 2003). Developmental flexibility is simply another process that alters the shape of the fitness surface.

The adaptive search problem has three phases, similar to the three aspects of search described in the prior section. First, partially matching expression to the environment brings the phenotype close to the new fitness landscape. Second, random perturbations of phenotype occur around the location set by the process of environmental matching. Third, any genotype on a novel fitness surface can survive and subsequently adapt by small genetic variations and natural selection.

Figure 9 illustrates the three aspects of adaptive search. Suppose a genotype expresses phenotypes centered at $\bar{x}$. In the first aspect of adaptive search, a genotype can adjust phenotypic expression to match the environment. The possible range of phenotypes varies from $\bar{x} - c$ to $\bar{x} + c$. The phenotype expressed by environmental matching is the new average value, around which random perturbations may occur. In the figure, the solid peak shows the fitness landscape imposed by a novel or extreme environmental challenge. The example genotype can come close to the new peak by modulating expression to produce an average phenotype of $\bar{x} + c$. However, if no random variation occurs around $\bar{x} + c$, that phenotype falls outside the range of phenotypic values that can survive.

In the second aspect of adaptive search, the genotype may produce phenotypes randomly distributed around the mean value of $\bar{x} + c$. Those random fluctuations smooth the fitness landscape, shown by the dashed curve. The average phenotype $\bar{x} + c$ can now survive. Matching the environment allowed expression of mean phenotype $\bar{x} + c$, and random fluctuations in phenotype smoothed the nearby landscape sufficiently.

Box 4: Vertebrate immunity

Invading pathogens present a vast diversity of foreign molecules that must be recognized. The vertebrate adaptive immune system develops antibodies by synergism between phenotypic and genetic variation, following the general three-part search process described in the text (Frank, 2002; Murphy *et al.*, 2007).

First, to generate genetic diversity, B cell lineages within the body undergo programmed genetic recombination early in life. That recombination yields genetically distinct cellular clones. Each clone produces a distinct antibody.

Second, each antibody type from this initial diversity tends to bind relatively weakly to a variety of foreign antigens. In this regard, the original or 'natural' antibodies trade the cost of weak binding for the benefit of a phenotypically diverse response – a broad reaction norm. Upon challenge with a foreign antigen, those B cells with matching antibodies are stimulated to expand clonally. That clonal expansion can be thought of as survival and reproduction of those genotypes that land on the fitness surface imposed by the unpredictable invader.

Third, the weakly binding antibodies undergo a programmed round of hypermutation to the antibody binding site and selection favoring variants that bind more tightly to the foreign antigen. This affinity maturation produces tightly binding and highly adapted antibodies in response to the novel challenge. Put another way, the initially stimulated antibodies on the edge of the 'fitness surface' climb the surface toward the fitness peak.

In the process of climbing the fitness peak by local genetic variation and natural selection, the refined antibodies match more closely to the environmental challenge. In particular, the refined antibodies narrow their reaction norm by increasing their binding affinity for close matches and reducing their binding affinity for slightly mismatched binding.

In summary, the ability of the adaptive immune system to respond to the huge diversity of potential challenges depends on its synergism between genetic variability and the reaction norm. The initial natural antibodies arise from genetically diverse clones produced by recombination. That genetic diversity by itself could not cover the huge space of possible challenges. The broad reaction norm around each genetic variant allows protection against novel challenge. Once partial recognition is achieved through the natural antibodies, the system refines the match locally by affinity maturation.

Once the genotype achieves survival, the third phase of adaptation may proceed. In this case, the mean phenotype $\bar{x} + c$ has low fitness on the dashed fitness surface. But the fitness surface has a smoothly increasing path to the peak of maximum fitness. Genetic variations in the genotype may shift the range of phenotypes that can be produced, allowing natural selection to drive the population up the fitness surface to the peak.

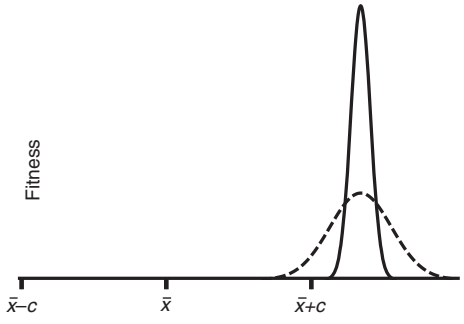

Fig. 9 Plasticity or learning provides a partial match of phenotype to novel or extreme environmental challenge. A genotype's default expression has average phenotypic value $\bar{x}$. An individual can modify average trait expression in response to the environment. The average expressed phenotype can be any value in the range $\bar{x} \pm c$. The environmental challenge defines the fitness landscape that relates phenotype to fitness, here shown as the solid peak. Phenotypic expression, modulated by a match to the environment, shifts the phenotype to $\bar{x} + c$. However, the fitness associated with a phenotype of $\bar{x} + c$ is zero, because that value remains outside the range of viable phenotypes. Suppose $\bar{x} + c$ is the average phenotype expressed, and random perturbations of expression cause variability in phenotype around that average value. The random component of phenotype smooths the fitness landscape, leading to the dashed fitness surface. The expressed average phenotype, $\bar{x} + c$, now falls within the smoothed fitness surface, allowing the genotype to survive. Subsequent adaptation may allow improved fitness, by altering the range of phenotypes that can be expressed so that a match to the fitness peak may be achieved.

Conclusion

Evolutionary theory emphasizes genetic variation as the source of evolutionary novelty. By the standard theory, the usual sequence would be a novel environmental challenge, genetic variation either already present or arising de novo, and evolutionary response to the novel environment by change in gene frequency.

In this classical evolutionary theory, genetics provides the source of phenotypic variation on which natural selection acts. By contrast, development may generate the phenotypic novelty that initiates adaptation to environmental challenge. The sequence would be novel environmental challenge, initial survival by those individuals with a phenotypic norm of reaction that overlaps the new fitness surface, and subsequent adaptation by genetic variants from those phenotypes that survive the initial challenge.

West-Eberhard (2003) traces the theoretical foundations of this topic from the late 19th century. Since that time, the idea that developmental processes may play a key role in initiating adaptation has never been popular. Evolutionary change is usually tied in thought to genetic change. Nonheritable phenotypic variation by itself is therefore usually believed not to accelerate evolutionary rate.

The original theories of learning, developmental plasticity, and reaction norms have always understood the relations between genotype, phenotype, environment,

Box 5: A simple prediction

Frequent exposure to novel or extreme environmental challenge favors a greater capacity for evolutionary response. Faster evolutionary response may be achieved by enhanced generation of genetic variability. For example, in experimental evolution studies, strong competition favors strains of bacteria with a high mutation rate (Loh *et al.*, 2010). Strong competition can be thought of as a form of extreme challenge. However, interference between genetically distinct competing clones complicates the association between mutation rate and evolutionary rate (Rainey, 1999).

Faster evolutionary response may also be achieved by enhanced phenotypic variability through learning or plasticity. However, it may be difficult to test empirically whether learning or plasticity are direct responses to increased environmental challenge.

The problem is that learning and plasticity tend to be complex adaptations. To build those complex adaptations takes a long time and many changes. That slow process means that one would have to measure the change in evolutionary pressure over the long period during which learning or plasticity evolve. Ultimately, one would need direct measures of environmental pressure in periods of weak challenge and in periods of strong challenge, and measures of the time course of change in learning or plasticity. That is hard to do.

A simpler prediction concerns the rapid increase in the reaction norm in response to an environmental challenge. If a population encounters a novel or extreme environmental challenge, it may respond by broadening its reaction norm. A broader reaction norm can be achieved simply by increasing the tendency for stochastic perturbations during development – a breakdown in the normal homeostatic processes that keep phenotypes within narrow bounds. This prediction about the broadening of the reaction norm follows immediately from Holland's (1975) contrast between exploitation and exploration and the quote given in the text by Haldane (1932). Many specific models analyze how the reaction norm responds to environmental challenge. Lande (2009) provided new models and a good overview of the literature.

In this case, one simply needs to measure the change in environmental pressure over short periods of time and the associated change in the reaction norm of the population. It would also be interesting to measure how a broader reaction norm acts synergistically with genetic variation to speed the evolutionary response.

and evolutionary change. However, the jargon from those theories is thick: the Baldwin effect, genetic assimilation, reaction norms, hopeful monsters, niche construction, and environmentally induced evolution. Each variant theory invoked special environmental conditions, developmental processes, and interactions with genetics. And each in its own way jousted with the ghost of Lamarck. A casual observer could be forgiven for steering clear of the whole mess. Wisdom suggested to wait for clear empirical examples. Induction still dominates mainstream thought in biology.

Many years ago, I read Hinton and Nowlan's (1987) article and Maynard Smith's (1987) related essay on the Baldwin effect. They focused on the essential theoretical point. Learning smooths the fitness surface, changing evolutionary dynamics in a way that greatly accelerates adaptation to novel or extreme environmental challenges. When one views the whole confusing field in that simple light, one sees that all the complexities of the theories and mechanistic details of phenotypic variability ultimately reduce to the same point. Developmental variation smooths the fitness landscape. A smoothed fitness landscape profoundly alters evolutionary dynamics, particularly in response to novel or extreme environmental challenge.

Acknowledgments

Marsha Rosner's insight into cellular variability and the evolutionary dynamics of cancer stimulated this work. My research is supported by National Science Foundation grant EF-0822399, National Institute of General Medical Sciences MIDAS Program grant U01-GM-76499, and a grant from the James S. McDonnell Foundation. I completed this article while supported by a Hogge-Baer Visiting Professorship in Cancer Research at the University of Chicago.

References

Altschuler, S.J. & Wu, L.F. 2010. Cellular heterogeneity: do differences make a difference? *Cell* **141**: 559–563.

Ancel, L.W. 2000. Undermining the Baldwin expediting effect: does phenotypic plasticity accelerate evolution? *Theor. Popul. Biol.* **58**: 307–319.

Anderson, R.W. 1995. Learning and evolution: a quantitative genetics approach. *J. Theor. Biol.* **175**: 89–101.

Aubret, F. & Shine, R. 2009. Genetic assimilation and the postcolonization erosion of phenotypic plasticity in island tiger snakes. *Curr. Biol.* **19**: 1932–1936.

Avery, S.V. 2006. Microbial cell individuality and the underlying sources of heterogeneity. *Nat. Rev. Microbiol.* **4**: 577–587.

Baldwin, J.M. 1896. A new factor in evolution. *Am. Nat.* **30**: 441–451.

Bell, A.M. & Robinson, G.E. 2011. Behavior and the dynamic genome. *Science* **332**: 1161–1162.

Booth, I.R. 2002. Stress and the single cell: intrapopulation diversity is a mechanism to ensure survival upon exposure to stress. *Int. J. Food Microbiol.* **78**: 19–30.

Chevin, L.M. & Lande, R. 2010. When do adaptive plasticity and genetic evolution prevent extinction of a density-regulated population? *Evolution* **64**: 1143–1150.

Chevin, L.M., Lande, R. & Mace, G.M. 2010. Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. *PLoS Biol.* **8**: e1000357.

Coyne, J.A., Barton, N.H. & Turelli, M. 1997. A critique of Sewall Wright's shifting balance theory of evolution. *Evolution* **51**: 643–671.

Crispo, E. & Rausher, M. 2007. The Baldwin effect and genetic assimilation: revisiting two mechanisms of evolutionary change mediated by phenotypic plasticity. *Evolution* **61**: 2469–2479.

De Jong, G. 1990. Quantitative genetics of reaction norms. *J. Evol. Biol.* **3**: 447–468.

Frank, S.A. 1996. The design of natural and artificial adaptive systems. In: *Adaptation* (M.R. Rose & G.V. Lauder, eds), pp. 451–505. Academic Press, San Diego, CA.

Frank, S.A. 2002. *Immunology and Evolution of Infectious Disease*. Princeton University Press, Princeton, NJ.

Frank, S.A. 2011. Natural selection. I. Variable environments and uncertain returns on investment. *J. Evol. Biol.* doi: 10.1111/j.1420-9101.2011.02378.x.

Gavrilets, S. 2004. *Fitness Landscapes and the Origin of Species*. Princeton University Press, Princeton, NJ.

Gavrilets, S. 2010. Rapid transition towards the division of labor via evolution of developmental plasticity. *PLoS Comput. Biol.* **6**: e1000805.

Gavrilets, S. & Scheiner, S.M. 1993. The genetics of phenotypic plasticity. V. Evolution of reaction norm shape. *J. Evol. Biol.* **6**: 31–48.

Geyer, C.J. & Thompson, E.A. 1995. Annealing markov chain monte carlo with applications to ancestral inference. *J. Am. Stat. Assoc.* **90**: 909–920.

Haldane, J.B.S. 1932. *The Causes of Evolution*. Cornell University Press, Ithaca, NY.

Hall, B. K. 2003. Baldwin and beyond: organic selection and genetic assimilation. In: *Evolution and Learning: The Baldwin Effect Reconsidered* (B.H. Weber & D.J. Depew, eds), pp. 141–167. MIT Press, Cambridge, MA.

Hinton, G.E. & Nowlan, S.J. 1987. How learning can guide evolution. *Complex Syst.* **1**: 495–502.

Holland, J.H. 1975. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI.

Kaneko, K. 2011. Characterization of stem cells and cancer cells on the basis of gene expression profile stability, plasticity, and robustness. *Bioessays* **33**: 403–413.

Kirkpatrick, S., Gelatt, C.D. & Vecchi, M.P. 1983. Optimization by simulated annealing. *Science* **220**: 671–680.

Kirschner, M. & Gerhart, J. 1998. Evolvability. *Proc. Natl Acad. Sci. USA* **95**: 8420–8427.

Lande, R. 1980. Genetic variation and phenotypic evolution during allopatric speciation. *Am. Nat.* **116**: 463–479.

Lande, R. 2009. Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *J. Evol. Biol.* **22**: 1435–1446.

Loh, E., Salk, J.J. & Loeb, L.A. 2010. Optimization of DNA polymerase mutation rates during bacterial evolution. *Proc. Natl Acad. Sci. USA* **107**: 1154–1159.

Mayley, G. 1997. Landscapes, learning costs, and genetic assimilation. *Evol. Comput.* **4**: 213–234.

Maynard Smith, J. 1987. When learning guides evolution. *Nature* **329**: 761–762.

Mayr, E. 1982. *The Growth of Biological Thought*. Harvard University Press, Cambridge, MA.

Mills, R. & Watson, R.A. 2006. On crossing fitness valleys with the Baldwin effect. In: *Proceedings of the Tenth International Conference on the Simulation and Synthesis of Living Systems* (L.M. Rocha, L.S. Yaeger, M.A. Bedau, D. Floreano, R.L. Goldstone & A. Vespignani, eds), pp. 493–499. MIT Press, Cambridge, MA.

Murphy, K.M., Travers, P. & Walport, M. 2007. *Janeway's Immunobiology*, 7th edn. Garland Science, Boca Raton, FL.

Niepel, M., Spencer, S.L. & Sorger, P.K. 2009. Non-genetic cell-to-cell variability and the consequences for pharmacology. *Curr. Opin. Chem. Biol.* **13**: 556–561.

Pigliucci, M. 2001. *Phenotypic Plasticity: Beyond Nature and Nurture*. Johns Hopkins University Press, Baltimore, MD.

Pigliucci, M. 2008. Is evolvability evolvable? *Nat. Rev. Genet.* **9**: 75–82.

Pigliucci, M., Murren, C.J. & Schlichting, C.D. 2006. Phenotypic plasticity and evolution by genetic assimilation. *J. Exp. Biol.* **209**: 2362–2367.

Price, T.D., Qvarnström, A. & Irwin, D.E. 2003. The role of phenotypic plasticity in driving genetic evolution. *Proc. R. Soc. Lond. B Biol. Sci.* **270**: 1433–1440.

Rainey, P.B. 1999. Evolutionary genetics: the economics of mutation. *Curr. Biol.* **9**: R371–R373.

Rajon, E. & Masel, J. 2011. Evolution of molecular error rates and the consequences for evolvability. *Proc. Natl Acad. Sci. USA* **108**: 1082–1087.

Rice, S.H. 2002. A general population genetic theory for the evolution of developmental interactions. *Proc. Natl Acad. Sci. USA* **99**: 15518–15523.

Rubin, H. 1990. The significance of biological heterogeneity. *Cancer and Metastasis Rev.* **9**: 1–20.

Schlichting, C.D. & Pigliucci, M. 1998. *Phenotypic Evolution: A Reaction Norm Perspective*. Sinauer Associates, Sunderland, MA.

Spencer, S.L., Gaudet, S., Albeck, J.G., Burke, J.M. & Sorger, P.K. 2009. Non-genetic origins of cell-to-cell variability in trail-induced apoptosis. *Nature* **459**: 428–432.

Sumner, E.R. & Avery, S.V. 2002. Phenotypic heterogeneity: differential stress resistance among individual cells of the yeast *Saccharomyces cerevisiae*. *Microbiology* **148**: 345–351.

Suzuki, R. & Arita, T. 2007. Repeated occurrences of the Baldwin effect can guide evolution on rugged fitness landscapes. In: *Proceedings of the First IEEE Symposium on Artificial Life (IEEE-ALIFE'07)*, pp. 8–14. IEEE Press, Hoboken, NJ.

Turney, P. 1996. Myths and legends of the Baldwin effect. In: *Proceedings of the Workshop on Evolutionary Computing and Machine Learning, at the 13th International Conference on Machine Learning (ICML-96)*, pp. 135–142. Bari, Italy.

Turney, P., Whitley, D. & Anderson, R.W. 1997. Evolution, learning, and instinct: 100 years of the Baldwin effect. *Evol. Comput.* **4**: iv–viii.

Waddington, C.H. 1942. Canalization of development and the inheritance of acquired characters. *Nature* **154**: 563–565.

Waddington, C.H. 1953. Genetic assimilation of an acquired character. *Evolution* **7**: 118–126.

Wagner, G.P. & Altenberg, L. 1996. Complex adaptations and the evolution of evolvability. *Evolution* **50**: 967–976.

West-Eberhard, M.J. 2003. *Developmental Plasticity and Evolution*. Oxford University Press, New York, NY.

Whitlock, M.C. 1997. Founder effects and peak shifts without genetic drift: adaptive peak shift occur easily when environments fluctuate slightly. *Evolution* **51**: 1044–1048.

Woods, R.J., Barrick, J.E., Cooper, T.F., Shrestha, U., Kauth, M.R. & Lenski, R.E. 2011. Second-order selection for evolvability in a large Escherichia coli population. *Science* **331**: 1433–1436.

Wright, S. 1931. Evolution in Mendelian populations. *Genetics* **16**: 97–159.

Yomo, T., Ito, Y., Sato, K. & Kaneko, K. 2005. Phenotypic fluctuation rendered by a single genotype and evolutionary rate. *Physica A* **350**: 1–5.

Received 13 June 2011; accepted 18 July 2011

JOURNAL OF Evolutionary Biology

doi: 10.1111/j.1420-9101.2011.02431.x

REVIEW

Natural selection. III. Selection versus transmission and the levels of selection*

S. A. FRANK

Department of Ecology and Evolutionary Biology, University of California, Irvine, CA, USA

Keywords:

genetic variation;
group selection;
kin selection;
life history;
mutation.

Abstract

George Williams defined an evolutionary unit as hereditary information for which the selection bias between competing units dominates the informational decay caused by imperfect transmission. In this article, I extend Williams' approach to show that the ratio of selection bias to transmission bias provides a unifying framework for diverse biological problems. Specific examples include Haldane and Lande's mutation–selection balance, Eigen's error threshold and quasispecies, Van Valen's clade selection, Price's multilevel formulation of group selection, Szathmáry and Demeter's evolutionary origin of primitive cells, Levin and Bull's short-sighted evolution of HIV virulence, Frank's timescale analysis of microbial metabolism and Maynard Smith and Szathmáry's major transitions in evolution. The insights from these diverse applications lead to a deeper understanding of kin selection, group selection, multilevel evolutionary analysis and the philosophical problems of evolutionary units and individuality.

> In evolutionary theory, a gene could be defined as any hereditary information for which there is a … selection bias equal to several or many times its rate of endogenous change (Williams, 1966, p. 44).

Introduction

Natural selection increases inherited information about environmental challenge. Against selection, imperfect transmission reduces inherited information. Many problems in biology come down to understanding the relative balance between selection and imperfect transmission.

A clear understanding of selection and transmission requires greater precision with regard to abstract notions such as *inherited information*. However, before heading off in pursuit of abstract theory, it pays to have some simple examples in mind. Those simple examples define the challenges for deeper theory.

In this paper, I work through several examples that turn on the relative strength of selection and imperfect transmission: Haldane (1927) and Lande's (1975) balance between selection and mutation, Eigen's (1992) error threshold and quasispecies, Van Valen's (1975) multilevel analysis of clade selection, Price's (1972) multilevel analysis of group selection, Szathmáry & Demeter's (1987) stochastic corrector model of early cellular evolution, Levin & Bull's (1994) short-sighted model of parasite evolution, Frank's (2010) timescale model of microbial metabolism and Maynard Smith & Szathmáry's (1995) major transitions in evolution.

Others have pointed out similarities between some of these examples (Maynard Smith & Szathmáry, 1995; Michod & Herron, 2006; Okasha, 2006). However, the broad unity with regard to selection and transmission is sometimes lost. In addition, the key role of timescale, although often noted, has not always been linked to selection and transmission in a simple and general way.

Williams' (1966) quote emphasizes timescale: the opposition between *selection bias* and *rate of endogenous change*. An entity can be shaped by natural selection only to the extent that the informational gain by natural selection is not overwhelmed by the relative rate of informational decay by imperfect transmission. The balance between selection and decay often turns on the relative timescales over which those forces operate.

Correspondence: Steven A. Frank, Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697–2525, USA.
Tel.: +1 949 824 2244; fax: +1 949 824 2181; e-mail: safrank@uci.edu
*Part of the Topics in Natural Selection series. See Box 1.

Box 1: Topics in the theory of natural selection

This article is part of a series on natural selection. Although the theory of natural selection is simple, it remains endlessly contentious and difficult to apply. My goal is to make more accessible the concepts that are so important, yet either mostly unknown or widely misunderstood. I write in a nontechnical style, showing the key equations and results rather than providing full derivations or discussions of mathematical problems. Boxes list technical issues and brief summaries of the literature.

The decay of transmission fidelity

Many processes reduce the similarity between ancestor and descendant. In classical genetics, mutation changes the intrinsic quality of an allele during transmission. Mixing of alleles reduces transmission fidelity because of interactions with the changed combination of other alleles. Internal selection changes the frequency of alleles within individuals, altering the similarity between ancestor and descendant.

Internal selection may occur within a pool of allele copies, in which certain alleles express traits that cause their frequency to increase against their neighbours (Burt & Trivers, 2008). For example, shortened mitochondrial genomes in certain yeast replicate faster than full genomes. The shortened genomes can rise in frequency within cells, even though they reduce individual-level fitness. In diploid Mendelian genetics, internal selection arises when traits increase allelic transmission to offspring to greater than the standard Mendelian probability of one-half.

Mutation or mixing of alleles may, in some cases, cause unbiased change during transmission. Unbiased change decays transmission fidelity but does not affect the direction of evolution for the average value of traits. Unbiased change can increase the variation in traits by causing random fluctuations in the characters expressed by descendants. Under stabilizing selection, the amount of variation may be shaped by a balance between an increase caused by fluctuations in transmission and a decrease caused by selection removing fluctuations from the favoured value (Lande, 1975).

Biased mutation or internal selection causes a directional change during transmission. When the directional change during transmission opposes selection between individuals or groups, the balance between selection and transmission influences the average value of traits.

Selection versus transmission

Total evolutionary change can be partitioned into components of selection and transmission:

$$\text{Total change} = \Delta\text{selection} + \Delta\text{transmission},$$

in which the symbol Δ means *the change caused by the process of* or *the change in the quantity of* depending on context. This partition of total change into selection and transmission is so important that it is worthwhile to express the partition with symbols. The symbolic form allows us to look at variations in the partition and the consequences for understanding evolutionary process (Box 2).

Total change can be expressed by the change in the average value of some trait. Let $\Delta\bar{z}$ be the change in the average trait value. Do not be misled by the word *average*. We can consider the average of the squared deviations of a trait to measure the variance, or the average of the product of different characters to measure correlations, or the average frequency of an allele in the population, or any other expression leading to some quantity: $\Delta\bar{z}$ is the change in whatever quantity we choose. We write total evolutionary change as $\bar{w}\Delta\bar{z}$, where $\bar{w}$ is average fitness. Average fitness accounts for the total numbers of births and deaths, allowing us to express selection and transmission directly in proportion to total change (see Box 2).

Express the change caused by selection as ΔS and the change caused by transmission as $\Delta\tau$. Then, the total change in symbols is

$$\bar{w}\Delta\bar{z} = \Delta S + \Delta\tau. \tag{1}$$

Any evolutionary problem can be expressed in this way. But whether it is useful to do so depends on the particular problem and, to some extent, on one's preference between alternative ways to partition total change into various components.

Selective improvement often pushes traits in the opposite direction from transmission decay. The balance between these opposing forces occurs when the total change is zero

$$\bar{w}\Delta\bar{z} = \Delta S + \Delta\tau = 0, \tag{2}$$

which also means that at an equilibrium balance

$$\Delta S = -\Delta\tau. \tag{3}$$

This equation provides the ultimate expression of a balance between selective improvement and transmission decay (Frank & Slatkin, 1990).

We can often write the change caused by selection as

$$\Delta S = s_z V_z, \tag{4}$$

where s_z is the selective intensity on the character z and V_z is the variance in the character z under selection (see Box 2). If selection causes a decrease in the character, we would instead write

$$\Delta S = -s_z V_z \tag{5}$$

to express the negative contribution of selection to the change in character. Using these expressions for the change caused by selection in eqn (3), we obtain the equilibrium variance under a balance between selection and transmission as

$$V_z = \left|\frac{\Delta\tau}{s_z}\right|. \tag{6}$$

The absolute value is used because s_z and V_z are always positive, whereas $\Delta\tau$ may be positive or negative depending on whether the transmission bias increases or decreases the trait. The key point is that when the opposing forces of selection and transmission are in balance, we have this simple expression for the variance of a character.

A measure of selection versus transmission

How exactly should we interpret Williams' phrase 'hereditary information for which there is a … selection bias equal to several or many times its rate of endogenous change'? We could evaluate the strength of selection bias relative to transmission bias to obtain a simple measure for the ratio, $\mathscr{R}$, between the forces. In particular, when the two terms oppose each other, we may write

$$\mathscr{R} = \log\left(-\frac{\Delta S}{\Delta\tau}\right). \tag{7}$$

The negative sign appears because the opposing directions of change for ΔS and $\Delta\tau$ mean that these terms have opposite signs. The negative sign makes the ratio positive. The logarithmic scaling imposes symmetry about zero. The ratio is zero when the two forces are equal, as in the balance condition of eqn (3). Increasingly positive values arise from greater dominance of selection bias, whereas increasingly negative values arise from greater dominance of transmission bias. Later examples illustrate the application of this ratio.

Multilevel selection

The individual typically comprises a group of alleles. In some cases, selection may occur between alleles within the individual. That selection within individuals creates a transmission bias between ancestors and descendants, because the sample of alleles transmitted to descendants is changed by selection between alleles within the ancestor. The total change can be expressed as selection between individuals plus the transmission bias created by selection within individuals. In this case, we can think of selection and transmission as the combination of two levels of selection (Price, 1972; Hamilton, 1975).

Now consider a population of individuals structured into groups. The total change may be partitioned into selection between groups and the transmission bias between an ancestral group and the descendants derived from that group (Box 3). Selection between individuals within the group will often strongly influence transmission bias, because selection within the group changes the composition of traits that are transmitted to descendants of that group. The total change can be expressed primarily as selection between groups and the transmission bias created by selection within groups. Once again, we can think of selection and transmission as the combination of two levels of selection.

Box 2: Price's selection and transmission

The Price equation provides a useful separation between selection and transmission (Price, 1970, 1972; Hamilton, 1975). Much literature and misunderstanding descend from the Price equation. I will treat the topic fully in a later article. Here, I briefly summarize the essential concepts. My previous publications related to the Price equation provide further background (Frank, 1995a, 1997b, 1998). Other key references lead into the broader literature (Wade, 1985; Heisler & Damuth, 1987; Michod, 1997a; Grafen, 2002; Rice, 2004; Okasha, 2006; Gardner, 2008).

I used the Price equation as the basis for eqn (1) in the text. The Price equation may be written as

$$\bar{w}\Delta\bar{z} = \mathrm{Cov}(w,z) + \mathrm{E}(w\Delta z).$$

Comparing with eqn (1), the selection bias is $\Delta S = \mathrm{Cov}(w,z)$. This simply says that the selection bias is the association between fitness and character value, where association is expressed by the covariance. The transmission bias is $\Delta\tau = \mathrm{E}(w\Delta z)$. This says that the transmission bias is the average (expectation) of the change in character value, Δz, between parent and offspring. The individual parent–offspring biases in transmission are weighted by parental fitness, w. If, for example, a parent reproduces little, then that parent's transmission bias contributes little to the average transmission bias in the population.

The expression for selection in eqn (4) is derived as $\Delta S = \mathrm{Cov}(w,z) = \beta_{wz}V_z$, because the covariance of w and z is the product of the regression coefficient, β, of w on z and the variance of z. Define $s_z = |\beta_{wz}|$, and apply a minus sign when $\beta_{wz} < 0$ to obtain eqn (5). See Frank (1997b) for the interpretation of these terms in the Price equation.

In the mutation–selection balance models, either $z \equiv q$ is allele frequency or z is the squared deviation of a trait from the optimum. In either case, z is always positive and the association between fitness and character value is negative. Thus, $-s_z = \beta_{wz}$, and we can express fitness in terms of the regression form

$$\mathrm{E}(w|z) = 1 + \beta_{wz}z = 1 - s_z z. \tag{20}$$

Here, I set maximum fitness to one. Any proportional change in maximum fitness is matched by the same proportional change in the regression coefficient, so the expression can be scaled arbitrarily. From this regression expression, the average of $s_z z$ must be less than one; otherwise average fitness drops below zero and mutational decay dominates selection, causing loss of heritable information or 'mutational meltdown' (Lynch *et al.*, 1993).

Note that the regression expression $\mathrm{E}(w|z) = 1 - \beta_{wz}z$ does not require a linear relation between character value and fitness. Rather, β_{wz} is simply the best least squares fit of fitness to trait value given the actual pattern by which trait values associate with fitness.

Box 3: What are groups?

One must distinguish between two aspects. On the one hand, the fundamental theory works perfectly for essentially any conception of groups of alleles, individuals or other entities. The groups do not require clear delineation, temporal continuity or biologically meaningful interaction. *Selection within groups* simply means the differential success between entities in the group, no matter how that differential success arises. *Transmission bias* simply means the fitness-weighted change in character value between the entities in the group and their descendants. No restriction is placed on how the descendants themselves are arranged into groups.

On the other hand, most potential groupings have no biological meaning. One naturally prefers groups defined by direct interaction, temporal continuity, shared interest and so on. Much literature debates alternative conceptions of meaningful groups (Maynard Smith, 1976; Wilson & Sober, 1989; Michod, 1997b; Gardner & Grafen, 2009). Difficulty occurs because the relative value of alternative views varies with biological context, intellectual goal and subjective bias about what is ultimately meaningful. Such undecidable alternatives attract endless debate and commentary.

Discussion of biologically meaningful alternatives can lead to improved understanding as the weight of evidence accumulates for certain views. However, that discussion has often sought absolute conclusions, when in fact context and subjective aspects necessarily play a role. In my view, one needs to keep in mind both the fundamental truth of the universal theory and the nuance of changing context and meaning in application. With both perspectives in mind, one never loses way.

Timescale

The balance between selection and transmission depends on the rate of selection between groups relative to the rate of endogenous change within groups. Timescale influences the relative rates.

Consider, for example, an increasing number of rounds of selection within groups for each round of selection between groups. If there is some limit to the ultimate size of groups, then the transmission bias caused by selection within groups increasingly dominates the selection between groups (Wilson & Colwell, 1981). Similarly, an increase in the number of rounds of replication within a lineage relative to the timescale of selection between lineages causes relatively greater mutation and decay during transmission compared with the selection bias. For example, the male mutation rate appears to be greater than the female mutation rate in several animal species, probably because of the greater number of replications per generation in the male germline (Nachman & Crowell, 2000).

It seems obvious that a relatively greater time for selection bias or transmission decay enhances the relative strength of a process. However, the simplicity of partitioning total change into selection and transmission in relation to timescale is not always developed clearly. By going through the examples properly, we can recover the simple conceptual unity that helps to explain a wide variety of biological problems.

Balance between selection and mutation

Perhaps the most basic of all evolutionary theory concerns the balance between selection and mutation (Haldane, 1927). From eqn (1), let the trait $\bar{z} \equiv q$ be the frequency of a deleterious allele. The equilibrium balance between selection and mutation occurs when the rate at which selection removes deleterious alleles equals the rate at which mutation adds new deleterious alleles. From eqn (3), the balance occurs when $\Delta S = -\Delta\tau$. From eqn (6), we can also express that balance as

$$V_q = \frac{\Delta\tau}{s_q}, \tag{8}$$

where V_q is the variance in allele frequency and s_q is the selective intensity on allele frequency. In this case, mutation increases the frequency of the mutant allele, so $\Delta\tau$ is positive and we do not need to use absolute values.

Classic results of population genetics

Suppose mutation changes a normal allele into a deleterious allele. Once an allele has become deleterious, it cannot mutate back into a normal allele. Let the mutation rate of normal alleles be μ. Normal alleles occur at frequency $1 - q$. Thus, the change in the number of mutant alleles caused by transmission bias is in proportion to $\Delta\tau = \mu(1 - q)$.

Selection reduces the reproductive success of mutant alleles by the selective intensity, $s_q \equiv s$. The variance in allele frequency is $V_q = q(1 - q)$, the standard binomial expression for variance when sampling a single allele. Substituting these expressions into eqn (8), the balance between selection and mutation occurs when the allele frequency is

$$q = \frac{\mu}{s}. \tag{9}$$

This result applies to haploid genetic systems and at least approximately to diploid systems with dominant deleterious mutations under the commonly used assumptions in population genetics. This expression captures the essential opposition between selective improvement and transmission decay that plays a key role in many biological problems. For the following

section, it will be useful to note that, from eqn (5), the selection bias is $\Delta S = -sV_q = -sq(1 - q)$.

To complete the classic treatment, I now write the case for a recessive mutation in diploid genetics. The mutation bias remains $\Delta\tau = \mu(1 - q)$. For recessive alleles, the deleterious phenotype is only expressed in the homozygote, which occurs at frequency q^2 under random mating. Thus, selective intensity on each copy of the deleterious allele increases with the probability, q, that it will be mated with another deleterious allele, so the selective intensity is $s_q = sq$. Substituting these expressions into eqn (8) yields the classic mutation–selection balance for recessive diploid genetics as

$$q = \sqrt{\frac{\mu}{s}}. \tag{10}$$

For the following section, it will be useful to note that the selection bias against the deleterious allele is $\Delta S = -sqV_q = -sq^2(1 - q)$.

Ratio of selection to transmission

The epigraph from Williams (1966) emphasizes the relative strength of selection bias to transmission bias. That comparison makes sense intuitively. However, when we use the results in this section to measure the relative strength of selection and transmission, the comparison turns out to be complex. The problem is that the relative strength of selection and transmission changes as evolution occurs in response to those forces.

For the simple models of selection and mutation in this section, Fig. 1 plots the relative strength of selection bias to transmission bias, $\mathscr{R}$, from eqn (7). For example, Fig. 1a shows the first model with equilibrium $q = \mu/s$ in eqn (9). In that case, $\Delta S = -sq(1 - q)$ and $\Delta\tau = \mu(1 - q)$, so the ratio is $\mathscr{R} = \log(sq/\mu)$.

The top curve of Fig. 1a plots $\mathscr{R}$ for $\log(s/\mu) = 5$. The plot scales the frequency of the mutant allele as $\log[q/(1 - q)]$. That scaling puts the midpoint of zero at $q = 0.5$, with high-frequency and low-frequency scaling symmetrically and roughly logarithmically about the midpoint.

In the top curve of Fig. 1a, when the mutant frequency is not too low, selection bias is many times the transmission bias, $\mathscr{R} \gg 0$. However, the mutant frequency evolves in response to the relative strength of selection and transmission. When selection is stronger, the mutant frequency, q, declines. As q declines, the ratio drops until $\mathscr{R} = 0$, at which point the selection bias equals the transmission bias. Similarly, when q is very small, the transmission bias is much greater than the selection bias, $\mathscr{R} \ll 0$, and the mutant frequency increases until the point $\mathscr{R} = 0$.

The ratio of the selection bias to the transmission bias does not have a constant value. As mutant frequency changes, the relative dominance of the two forces shifts. The system comes to rest only when selection and transmission are in balance. Given the changing relation between selection and transmission, how should we interpret Williams' dictum?

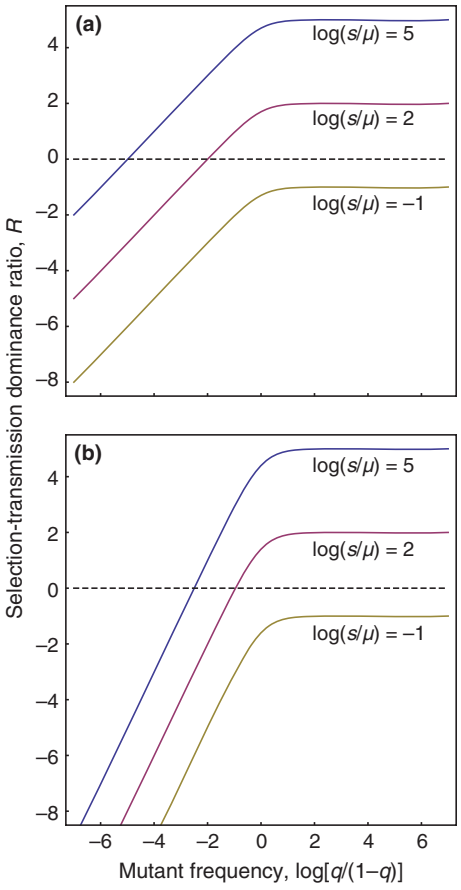

Fig. 1 The relative dominance of selection bias versus transmission bias in the models of selection and mutation. Relative dominance is measured by the ratio, $\mathscr{R}$, of eqn (7). (a) The diploid dominant or haploid model. (b) The diploid recessive model, in which $\Delta S = -sq^2(1 - q)$ and $\mathscr{R} = \log(sq^2/\mu)$. All logarithms use base 10.

We could emphasize the example of the lower curve in Fig. 1a. That curve never rises above zero, because transmission bias is always greater than selection bias for all frequencies. In that case, no hereditary information accumulates. So we might say that hereditary information accumulates when selection bias is stronger than transmission bias for at least some conditions. But that is a rather weak statement, changing Williams apparently beautiful clarity into a muddle.

Let us hold the point for now. As we go through various examples, we will see that the ratio of selection bias to transmission bias changes in response to key aspects of the particular problem under study. Rather than trying to abstract away how each particular problem shapes the changing ratio between selection and transmission, it may be more useful to use that ratio to

understand each particular problem and the relations between different problems.

Timescale

Timescale arises implicitly in these models, because selection and transmission are both expressed per unit time. In the simplest models, one usually considers a single round of mutation per generation for each round of selection per generation. However, multiple rounds of mutation can occur for each round of selection. For example, many replications typically occur in the male germline of species that make large numbers of sperm. Those multiple replications occur for each round of selection. The multiple replications apparently increase the mutation rate in relation to the strength of selection (Nachman & Crowell, 2000). This change in the relative magnitudes of selection and mutation is important but not particularly profound. Later, we will see more interesting ways in which timescale alters the balance between selection bias and transmission bias.

Variance under a balance between mutation and stabilizing selection

Selection sometimes acts in a stabilizing way, pushing the average phenotype towards an intermediate optimum. Mutation opposes selection by spreading trait values and increasing the average distance from the optimum. The decrease in phenotypic variance caused by selection is opposed by the increase in variance caused by mutation.

Here, I assume that all phenotypic variance is caused by simple genetics. This assumption allows me to focus on the processes that balance selection and mutation. I summarize the standard approach for this problem (Lande, 1975; Turelli, 1984; Barton & Turelli, 1987), following Frank & Slatkin (1990).

General expressions

Define the character of interest as $\gamma = z^2$, and set the optimum at zero, which is also the average value in this symmetric model. Then, z^2 is the squared distance from the optimum, and the average of this squared distance is the variance. Using γ as the character of interest, at mutation–selection balance, from eqns (6) and (8), we have

$$V_\gamma = \frac{\Delta\tau}{s_\gamma}. \tag{11}$$

Suppose a mutation adds or subtracts c from the phenotypic value, z. The two directions of change occur with equal probability. Thus, each mutation changes phenotype by $\pm c$. The contribution of each mutation to the change in squared deviation of phenotype, z^2, is, on average, c^2. Mutations happen at a rate μ, so the change in the phenotypic variance caused by mutation is

$$\Delta\tau = c^2\mu.$$

The scaling c^2 translates between genetic mutations and phenotypic effects. We can use that same scaling to translate between the phenotypic scale, γ, and the genetic scale, α, with the relation $\gamma = c^2\alpha$. Here, α is the squared deviation on the genetic scale, and γ is the squared deviation on the phenotypic scale. The average of squared deviations is the variance, so we have $\bar{\gamma}$ and $\bar{\alpha}$ for the phenotypic and genetic variances, where the overbar denotes the average.

The term V_γ is the variance of the squared phenotypic deviations, γ. Because a variance is itself a squared value, V_γ summarizes the square of the squared deviations, thus scaled to the fourth power. Therefore, the proper relation to go from the phenotypic scale to the genetic scale is $V_\gamma = c^4 V_\alpha$.

Substituting $V_\gamma = c^4 V_\alpha$ and $\Delta\tau = c^2\mu$ into eqn (11) yields

$$V_\alpha = \frac{\mu}{s}, \tag{12}$$

where $s = c^2 s_\gamma$. This expression for V_α provides the most general solution for variation under a balance between mutation and stabilizing selection. However, V_α is the variance of squared deviations

$$V_\alpha = \overline{\alpha^2} - \bar{\alpha}^2,$$

and thus scales with the fourth power of deviations. Typically, we seek expressions for the variance under stabilizing selection rather than expressions scaled to the fourth power of deviations. We can, under two particular cases, reduce the fourth power expression to an expression for variance under stabilizing selection.

Equilibrium variance

When selection is much stronger than mutation, $s \gg \mu$, the general balance result of eqn (12) is approximately

$$\bar{\alpha} \approx \frac{\mu}{s}, \tag{13}$$

where $\bar{\alpha}$ is the variance on the genetic scale. Note that this result is essentially the same as the haploid mutation–selection balance result in eqn (9) from the previous section. Box 4 provides the derivation.

When selection is much weaker than mutation, $s \ll \mu$,

$$\bar{\alpha} \approx \sqrt{\frac{\hat{\mu}}{s}}, \tag{14}$$

which matches the result for the diploid recessive model in eqn (10). Here, $\hat{\mu} = \mu/2$. With weak selection, most alleles deviate from the optimum of zero. At nonzero values, mutation is equally likely to move the allelic

Box 4: Variance under mutation–selection balance

To obtain the equilibrium genetic variance in eqn (13) when selection is much stronger than mutation, $s \gg \mu$, note that $\gamma = c^2\alpha$. Thus, with strong selection, most alleles will be at the optimum with $\gamma = \alpha = 0$, and a few alleles will be one mutational step away from the optimum at $\gamma = c^2$ and $\alpha = 1$ (Frank & Slatkin, 1990). Let the mutant frequency be q, so that $\alpha = 1$ with probability q, therefore $\alpha^2 = 1$ with probability q. Thus, $\bar{\alpha} = \overline{\alpha^2} = q$ and $\bar{\alpha}^2 = q^2$. With small q, we have $q \gg q^2$, therefore $V_\alpha = \overline{\alpha^2} - \bar{\alpha}^2 \approx \overline{\alpha^2} = \bar{\alpha}$, and thus eqn (12) leads to eqn (13).

To obtain the equilibrium genetic variance in eqn (14) when selection is much weaker than mutation, $s \ll \mu$, we assume that the distribution of allelic values approximately follows a Gaussian with a mean at zero (Kimura, 1965; Lande, 1975; Frank & Slatkin, 1990). With a Gaussian, the fourth moment is approximately three times the square of the second moment (variance), and thus $\overline{\alpha^2} \approx 3\bar{\alpha}^2$ and $V_\alpha = \overline{\alpha^2} - \bar{\alpha}^2 \approx 2\bar{\alpha}^2$. Using this expression for V_α in eqn (12) yields eqn (14).

value closer or farther from the optimum. Thus, only one-half of mutations are deleterious, and $\hat{\mu}$ expresses the deleterious mutation rate. Box 4 provides the derivation.

Note that selection on phenotypes can be strong, yet the selection bias against each mutational step can be weak. Here, *weak selection* refers to the effect on each mutational step. In particular, I defined $s = c^2 s_\gamma$ below eqn (12). If the phenotypic effect, c, of each mutation is small, then strong selection on the phenotypic scale, s_γ, can be associated with weak selection on each mutational step of size c when expressed on the genetic scale, s.

Mutation overwhelms selection

If the decay in fitness by mutation exceeds the maximum fitness that can be achieved, then mutation overwhelms selection. Mutation dominates selection when the magnitude of mutational effects is much greater than the magnitude of selection, $s \ll \mu$, which corresponds to results above for weak selection.

From eqn (20) of Box 2, we can write fitness as $w = 1 - s\alpha$, using $s \equiv s_\alpha$ for selective intensity on the genetic character α. Thus, average fitness is $\bar{w} = 1 - s\bar{\alpha}$ and, using eqn (14) for $\bar{\alpha}$, we obtain $\bar{w} = 1 - s\sqrt{\hat{\mu}/s} = 1 - \sqrt{\hat{\mu}s}$. Mutational meltdown occurs when $\bar{w} < 0$, which implies $\hat{\mu}s > 1$.

This condition simply means that the amount of deleterious mutation, $\hat{\mu}$, scaled by the fitness consequence per mutation, s, reduces fitness by an amount that is greater than maximal fitness. The next section considers when the mutation rate might be so high.

Error threshold and quasispecies

Eigen applied the fundamental tension between mutation and selection to the evolution of nucleotide sequences. In early evolution, the mutation rate was likely to be high because enzymes that correct replication errors did not yet exist. The initially high mutation rate and lack of error correction lead to Eigen's error threshold paradox (Eigen, 1971, 1992; Eigen & Schuster, 1977; Maynard Smith, 1979).

Suppose the initial replicating sequences had a length of n nucleotides. If the mutation rate per nucleotide is μ, then the mutation rate per sequence is roughly $n\mu$. The deleterious effect per mutation is s. Thus, the expected deleterious effect of mutation during each replication of a sequence of length n is $n\mu s$. When the deleterious effect per replication is greater than maximum fitness, here scaled to be one, mutation overwhelms selection and no selective increase in adaptation can be achieved. The condition for remaining below this error threshold is $n\mu s < 1$, which means that sequence length is limited to

$$n < \frac{1}{\mu s}.$$

Eigen noted the paradox of the error threshold for early evolution. Without error-correcting enzymes, the mutation rate was high. A high mutation rate limited the maximum sequence length. A short sequence could not contain enough information to encode error-correcting enzymes. Without error-correcting enzymes, the sequence remains too short to encode error correction. How did the biochemical machinery of error correction evolve?

Eigen *et al.* (1988, 1989) discussed a second interesting property of sequence evolution under mutation and selection. A population of sequences exists as a mixture of the most fit sequence and a variety of mutant sequences. Eigen called the most fit sequence the *master sequence*, and the population of sequences that are zero, one, two or more mutational steps away from the master sequence the *quasispecies*. The term *quasispecies* is meant to differentiate a population of variants from a typological notion of a species as a fixed, nonvarying entity.

The error threshold and the quasispecies are equivalent to the standard evolutionary concepts of heritable variation maintained by a balance between mutation and selection, as described in the previous section (Wilke, 2005). The epigraph from Williams captured the key idea of the error threshold by expressing the notion of a gene 'as any hereditary information for which there is a … selection bias equal to several or many times its rate of endogenous change' (Williams, 1966, p. 44). The classical mutation–selection theory of Haldane, extended to the maintenance of variation under stabilizing selection, expresses the concept of quasispecies. All of these theories have to do with the fundamental partition of total evolutionary change into a component of selection and a component of transmission fidelity.

Multilevel analysis of clade selection

Williams (1992) argued that the relative rates of selection and transmission influence evolutionary change at all taxonomic levels. Williams adopted the term *clade selection* from Stearns (1986). Van Valen's (1975) analysis provides the clearest way to understand the ideas and potential importance.

Van Valen (1975) began by comparing the evolutionary history of sexual and asexual types. He set up the problem by assuming that asexuals have a short-term advantage in growth rate relative to sexuals and that sexuals have a long-term advantage with regard to avoiding extinction and forming new species (Fisher, 1930; Stebbins, 1950). With those assumptions, Van Valen (1975, p. 87) suggests that one

> Consider a large set of species, some obligatory apomicts [asexuals] and some at least facultatively sexual. The apomicts will have a greater probability of extinction of lineages and the sexual species will have a greater probability of speciation by splitting of lineages. … However, apomicts will sometimes originate from sexual species because of their immediate advantage.

Van Valen recognized two levels of selection. Clades with more sexual species will increase in species number relative to clades with fewer sexual species. Thus, sex has an advantage between clades. Within clades, asexuals will arise repeatedly because of their short-term advantage relative to their sexual ancestor. The selection within clades that favours asexuals can be thought of as a transmission bias: sexual species sometimes produce asexual descendants, whereas asexual species rarely produce sexual descendants.

Van Valen used the fact that one can express the two levels of selection as selection between clades and a transmission bias within clades to develop a simple model for the equilibrium frequency of asexuals. That equilibrium frequency balances selection bias between clades favouring sexuals, with rate s, against transmission bias within clades favouring asexuals, with rate μ, to obtain the approximate equilibrium frequency of asexuals, q, as

$$q \approx \frac{\mu}{s}.$$

This expression is the same as the standard model of mutation–selection balance in genetics given in eqn (9). In this model, Van Valen (1975) emphasizes that selection at any taxonomic level is always potentially balanced against the rate of endogenous change at that level, echoing the epigraph from Williams. Endogenous change may arise in various ways, such as mutation by change of state or selection between the lower-level entities that comprise the higher level.

Van Valen also applied this approach to mammals. In mammals, genera with larger body size survive longer than genera with smaller body size, but the smaller-bodied genera bud off new genera at a higher rate. The net reproductive rate of small genera is higher, giving a selective advantage to small-bodied genera over large-bodied genera. Within genera, there is a bias towards larger body size. The distribution of mammalian body size is influenced by the balance between selection between genera favouring smaller size and selection within genera favouring larger size.

Various philosophical issues in the interpretation of clades as units have been taken up by Van Valen (1988), Williams (1992) and Okasha (2006). Here, I only applied the fact that one can partition the patterns of change at one level, such as clades, into components of selection and transmission. The philosophical issues focus on whether one can think of clades as natural units, for some reasonable meaning of *natural*.

Multilevel analysis of kin and group selection

Total evolutionary change includes a part caused by selection and a part caused by the lack of fidelity in transmission (eqn 1). In this section, I use that basic partition of total change to study two levels of selection, generalizing the model of clade selection in the prior section.

At the higher level, a group may be any sort of collection. We may, for example, consider groups of individuals or groups of alleles within an individual. Selection concerns differential success among groups. At the lower level, selection within groups causes a bias in the transmission fidelity of group-level characteristics (Lewontin, 1970; Price, 1972; Hamilton, 1975; Wilson & Sober, 1989; Okasha, 2006).

The most interesting problems arise when selection among groups opposes the transmission bias caused by selection within groups. We may then consider a balance between selection and transmission or, equivalently, a balance between the two levels of selection, $\Delta S = -\Delta\tau$, as in eqn (3).

I present three aspects of multilevel selection. First, I write a very simple expression for the balance between the two levels of selection. This expression of balance provides the general basis for multilevel models of selection and the analogy to the classical models of selection and mutation.

Second, I apply the balance between different levels of selection to the tension between competition and cooperation. That simple model illustrates how easily we can understand the basic processes of group-level cooperation within the broader framework of selection and transmission. I also show the fundamental equivalence of group selection and kin selection models in group-structured populations.

Third, I discuss the roles of population regulation and timescale. For population regulation, if limited space or resources regulate group productivity, then all groups may have roughly the same reproductive output. In that case, little selection occurs among groups, and the

within-group component of selection dominates (Wade, 1985; Frank, 1986, 1998; Taylor, 1992; Wilson *et al.*, 1992; Queller, 1994). For timescale, the number of rounds of selection within groups relative to the rate of selection among groups sets the relative scaling of selection between the two levels. When the rate of selection within groups overwhelms the rate of selection among groups, the within-group component of selection dominates evolutionary process (Williams, 1966).

The balance between levels of selection

The fundamental equation for balance is $\Delta S = -\Delta\tau$ the balance between selection bias and transmission bias. For multilevel selection, we interpret ΔS as the selection among groups and $\Delta\tau$ as the transmission bias caused by selection within groups. For problems in which selection at the different levels pushes character values in opposing directions, we may rewrite the expression as $\Delta S_a = -\Delta S_w$, the balance between selection among groups and selection within groups.

The change in a character caused by selection can be expressed as $\Delta S = sV$, the product of the selective intensity, s, and the variance in the character under selection, V (see Box 2). Thus, we may write the balance $\Delta S_a = -\Delta S_w$ as

$$s_a V_a = -s_w V_w.$$

In a group-structured population, the total variance is the sum of the variance among groups and the variance within groups, which we express as $V_t = V_a + V_w$. Making the substitution $V_w = V_t - V_a$ yields

$$s_a V_a = -s_w (V_t - V_a).$$

It is convenient to express the pattern of variance by the correlation coefficient $r = V_a/V_t$, where r measures the correlation in character values between individuals within a group. Dividing both sides by V_t yields

$$s_a r = -s_w (1 - r). \tag{15}$$

To understand this expression, we need to consider the interpretation of the correlation, $r = V_a/V_t$. The correlation is the fraction of the total variance that is among groups. Because variance provides a weighting on selection, r can be thought of as the fraction of the total weighting of selection that happens at the group level, and $1 - r$ can be thought of as the fraction of the total weighting of selection that happens within groups.

Thus, $s_a r$ is the intensity of selection among groups, s_a, multiplied by the weighting of selection at the group level, r. At a balance, the group-level component must be equal and opposite to the intensity of selection within groups, s_w, multiplied by the weighting of selection within groups, $1 - r$.

The correlation r is also a particular form of the regression coefficient of relatedness from kin selection theory, as hinted initially by Hamilton (1975, 1979) following from the work of Price (1972) and later analysed more formally (Grafen, 1984; Wade, 1985; Frank, 1986, 1998; West *et al.*, 2007). The equivalence of r and Hamilton's formal theory of kin selection establishes the exact equivalence of multilevel group selection and kin selection.

The tension between competition and cooperation

We need an explicit expression for the relation between a trait and fitness in order to evaluate the abstract expressions from the previous section. In this section, I present a simple model of competition and cooperation (Frank, 1994, 1995b).

In a group-structured population, we can express fitness as the product of two components. The first component is the individual's relative share of total group success. The second component is the total success of the group. For the first component, we may write the individual's relative share of the group's success as z/z_g, where z is the individual's tendency to be competitive against neighbouring group members for access to local resources, and z_g is the average competitive tendency in the individual's group. Selection within groups always favours greater competitive tendency, because an individual's share of group success always rises with an increase in z.

For the second component, total group success, suppose that the more intensely individuals compete against neighbours, the less efficient the group is in using its resources productively. For example, a certain fraction of local energy may go into outcompeting neighbours rather than enhancing productivity. We may express the negative effect of competitiveness on group productivity by writing the total group productivity as $1 - z_g$, in which the total productivity declines as the group members' average tendency to compete, z_g, rises. Thus, selection among groups always favours a less competitive and more cooperative behavioural tendency, because group success declines as average competitiveness, z_g, rises.

Putting the two pieces together, the fitness, w, of an individual with competitive tendency, z, in a group with average competitive tendency, z_g, is

$$w = \frac{z}{z_g}(1 - z_g). \tag{16}$$

To evaluate the balance between selection at the group level and selection within groups, we need to relate the expression for fitness to the particular selective tendencies and variance components in eqn (15). The selective intensity among groups is $s_a = -1$, because group fitness is $1 - z_g$, and selective intensity is the change (partial derivative) in group fitness with change in the average

trait value in the group. The selective intensity within groups is $s_w = (1 - z_g)/z_g$, which is the change in individual fitness, w, with change in individual character value, z.

Substituting these values for s_a and s_w into eqn (15) yields a balance between group and individual selection when

$$-r = -\frac{1 - z_g}{z_g}(1 - r). \tag{17}$$

Skipping over the technical details, we may say roughly that, in this case, selection acts in a stabilizing way, causing individuals' trait values to converge towards a single value that is an evolutionarily stable strategy (ESS). Thus, individual values, z, converge towards group averages, z_g, which in turn converge to a global value, z^*. Making the substitution $z_g = z^*$ and solving for z^* give the balance point (Frank, 1994, 1995b) as

$$z^* = 1 - r. \tag{18}$$

This balance point expresses the key insights into multilevel selection and kin selection. In terms of multilevel selection, $1 - r$ is the fraction of the total variance that occurs within groups. The greater this weighting of within-group selection, the higher the balancing point of z^*, the tendency of individuals to compete with neighbours. As variance shifts towards the group level, $1 - r$ declines, z^* decreases because competitive restraint is more strongly favoured, and the balance of selective forces increasingly favours cooperative behaviour. In terms of kin selection, as the coefficient of relatedness, r, increases, competitive restraint and cooperative behaviour rise.

Population regulation and timescale

Several factors may influence the intensity of selection within groups compared with the intensity of selection among groups (Alexander & Borgia, 1978; Wade, 1985). For example, if limited space or resources regulate group productivity, then all groups may have roughly the same reproductive output. In that case, little selection occurs among groups, and the within-group component of selection dominates (Wade, 1985; Frank, 1986, 1998; Taylor, 1992; Wilson *et al.*, 1992; Queller, 1994)

In the model from the prior section, group productivity was $1 - z_g$, and the change in group productivity with a change in the group phenotype, z_g, was $s_a = -1$. Suppose instead that the relation between group phenotype and group productivity is much weaker, because extrinsic aspects of space and resources limit group productivity. For example, if group productivity is $1 - \epsilon z_g$, where $\epsilon < 1$, then $s_a = -\epsilon$. Using that value of s_a in eqn (17), we obtain the solution

$$z^* = \frac{1 - r}{1 - r(1 - \epsilon)}$$

This solution is equivalent to eqn (18) when $\epsilon = 1$. As limits on group productivity become more stringent, ϵ declines towards zero, the balance tips more strongly in favour of selection within groups, the level of competitiveness, z^*, increases and, equivalently, the level of cooperation declines. There are, of course, many complex ways in which individual traits may relate to group productivity and to the intensity of selection within groups. But all the different complexities tend to reduce to the simple balancing of forces between selection among groups versus bias in transmission fidelity of group characteristics or, equivalently, selection among groups versus selection within groups. If some additional force weakens selection among groups, then selection within groups increasingly dominates. Similarly, if some additional force weakens selection within groups, then selection among groups increasingly dominates.

Timescale provides another example. If the number of rounds of selection within groups increases relative to the pace of selection among groups, then selection within groups increasingly dominates the balance of forces (Frank, 1986, 1987). I discuss two particular cases in later sections on parasite virulence and microbial metabolism.

Ratio of selection at different levels

In multilevel selection, the relative strength of selective bias to transmission bias from eqn (7) compares selection among groups with selection within groups. Substituting the expressions for ΔS and $\Delta\tau$ derived from eqn (16) into eqn (7) yields

$$\mathscr{R} = \log\left(\frac{z}{1 - z}\right) + \log\left(\frac{r}{1 - r}\right). \tag{19}$$

Figure 2 plots $\mathscr{R}$ versus the level of competitiveness, z, for different levels of relatedness, r.

The level of competitiveness is in equilibrium balance, z^*, when the lines cross $\mathscr{R} = 0$. At that point, selection bias among groups is equal and opposite to transmission bias caused by selection within groups. Once again, we see that selective bias is greater than transmission bias only when the system is out of equilibrium.

Following the epigraph from Williams, one may wish to think of groups as acquiring information, adaptation or a degree of unitary function to the extent that selective bias tends to dominate transmission bias. Because relative dominance depends on the phenotype, z, one interpretation would be that significant group-level function requires the relative dominance of selection over transmission across a wide range of possible phenotypes (Gardner & Grafen, 2009). The range of phenotypes over which selection bias dominates transmission bias increases with a rise in relatedness, r. Thus, one may say that increasing relatedness shifts the locus of information or adaptation towards the higher level.

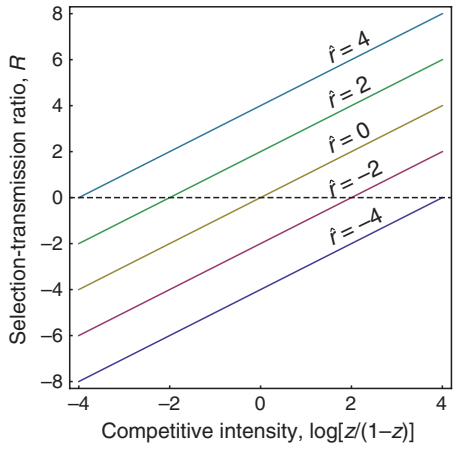

Fig. 2 The relative dominance of selection bias among groups versus transmission bias within groups in a multilevel selection model. Relative dominance is measured by the ratio, $\mathscr{R}$, of eqn (19). Different levels of relatedness shift the balance between selection bias among groups and transmission bias within groups. Here, relatedness is measured by $\hat{r} = \log[r/(1 - r)]$.

That interpretation of group-level unity or adaptation goes beyond what the analysis by itself presents. The analysis simply describes the way in which $\mathscr{R}$ shifts with competitive intensity and relatedness. The interpretation of group-level unity is a gloss that may aid or hinder understanding in different contexts. Ultimately, one must retain a clear view of the underlying analytical basis.

Stochastic corrector model of early protocells

Protocells are simple membrane-bound groups of genes that likely formed in early evolution (Maynard Smith & Szathmáry, 1995). A model of protocell evolution provides insight into group selection, kin selection, parasite virulence and the evolution of symbionts (Szathmáry & Demeter, 1987; Frank, 1994, 1996c, 1997a).

In the protocell model, the selective bias between cells opposes the transmission bias arising from selection between genes within cells. Expanding on the epigraph from Williams (1966), the degree to which adaptive design occurs at the protocell level versus the internal genic level depends on the selective bias between cells relative to the rate of endogenous change within cells.

Each protocell can be thought of as a bag that starts with k pieces of genetic material (chromosomes). The chromosomes compete within the protocell for resources. Success at acquiring resources influences the rate at which chromosomes can replicate themselves within the cell. More competitive chromosomes use up local resources less efficiently and reduce the overall success of the protocell and its group of chromosomes.

A protocell competes with other protocells for resources from the environment. A protocell produces a progeny cell after it has acquired sufficient resources and its chromosomes have replicated. The fitness of the protocell and its chromosomes depends on the rate of progeny production. Sampling of chromosomes occurs when progeny are formed: k chromosomes are chosen randomly from the pool of copies in the cell. I refer to this sampling process as segregation.

This protocell model is a particular expression of the group selection model in the previous section. By studying this particular example, we can see more clearly how specific aspects of mutation, competition and selection within groups affect transmission bias.

Suppose that the fitness of a chromosome follows the expression in eqn (16), repeated here

$$w = \frac{z}{z_g}(1 - z_g),$$

where z is a chromosome's tendency to be competitive against neighbouring chromosomes for access to local resources within the protocell, and z_g is the average competitive tendency of chromosomes in the protocell. Following eqn (18) of the previous section, the balance of selection between protocells and transmission bias within protocells is $z^* = 1 - r$, where r is the kin selection coefficient of relatedness among the chromosomes within a cell.

Virulence and symbiosis

The stochastic corrector model allows us to connect the abstract expressions from the multilevel analysis of kin and group selection to specific interpretations of parasite virulence and the evolution of symbionts within hosts (Frank, 1994, 1996c). For virulence, one can think of each of the k chromosomes as a parasite, and one can think of the protocell as the host. Competition between the parasites may cause inefficient use of host resources. Overexploitation of the host reduces host fitness. Thus, competition between parasites within hosts tends to increase virulence. The lower the relatedness, r, among the parasites within a host, the greater the competitiveness and virulence of the parasites, $z^* = 1 - r$.

We may also think of the k chromosomes as symbionts living within a host. From the host's point of view, increasing r reduces the competitiveness between the symbionts, aligning symbiont and host interests. In order to increase r, hosts may be favoured to reduce the number, k, of symbionts transmitted to offspring or transmitted between hosts (Frank, 1996a). Hosts may also be favoured to reduce the mixing of symbionts between different hosts (Frank, 1996b).

Kin selection and group selection

This model allows us to evaluate the meaning of the kin selection coefficient, r, within a particular scenario.

Assume that transmission is purely vertical, because the chromosomes do not mix between cells. In this model of vertical transmission, three forces affect the evolution of competitiveness, $z^* = 1 - r$.

First, selection between protocells favours reduced competitiveness of chromosomes within cells, leading to greater efficiency at the cellular level. Against that cellular level effect, the competition and selection between chromosomes within cells cause a transmission bias that favours increased competitiveness of chromosomes.

Second, mutations reduce the similarity among chromosomes within hosts, thus reducing r. The force imposed by mutation is controlled by two parameters, the mutation rate, μ, and the change in character value caused by each mutation, δ. Each mutational event changes z by $\pm\,\delta$, where the alternative directions of change occur with equal probability. Thus, mutation by itself causes no transmission bias.

Third, segregation samples from the local chromosomes when the protocell reproduces. Each new progeny starts with k chromosomal copies. When the cell reproduces, replicates of the local chromosomes are chosen stochastically according to the relative fitness within the cell, z/z_g. This sampling reduces the variance within hosts and increases relatedness.

A stochastic computer simulation of this model showed that relatedness, r, and equilibrium trait values, z^*, are held in balance by a delicate interaction between mutation, selection and segregation (Frank, 1994). The observed equilibrium trait values in the computer simulation closely follow the prediction $z^* = 1 - r$, where $r = V_a/V_t$ is calculated directly from the simulation by measuring the within-cell and total variances of trait values for the individual chromosomes in the population. The specific parameters affect variances and equilibrium trait values as expected: relatedness declines and z^* rises as the mutation rate, μ, or mutation step, δ, increases. An increase in the number of chromosomes per cell, k, causes an increase in competitiveness, z, because more copies reduce the variation among cells caused by sampling during segregation.

Kin selection arises from patterns of variance, not genealogy

The analysis in the previous section demonstrates that genealogy does not provide a sufficient explanation for the evolution of cooperative and competitive traits. The genealogical closeness between chromosomes in a cell increases as k declines. That genealogical aspect explains some of the changes in competitiveness, z^*. But, for a fixed genealogical scheme and a fixed mutation rate, the magnitude of the effect of each mutation, δ, can strongly influence the equilibrium value, z^* (Frank, 1994). Larger mutational effects raise the variance within groups relative to the variance among groups, causing a decline in r, an increase in the strength of selection within cells and an increase in the equilibrium competitiveness, z^*.

The theory of kin selection formulated by Hamilton (1970) depends solely on the patterns of variance and correlation, not on genealogy (Frank, 1998). Genealogy is often closely associated with the patterns of variance and correlation. The simple protocell model illustrates how the association between genealogy and the patterns of variance and correlation may break down. When the association breaks down, the true causal processes of variance and correlation explain the outcome. Since Hamilton's (1970) work, no fundamentally derived theory of kin selection based on genealogy has existed. However, it is often convenient to use the fact that genealogy typically associates with the underlying causal processes of variance and correlation. That convenience has unfortunately confused many authors about the distinction between a convenient association and the fundamental theory and its history.

We may recover the association between genealogy and causal process if mixing of chromosomes between cells occurs. Such mixing often dominates mutation in determining the patterns of variance within and among groups. In that case, genealogy may become the main force determining r and the equilibrium level of competitiveness, z^*.

In conclusion, the mutation rate and the size of mutational effects primarily influence the patterns of variance under some conditions, whereas the migration rate and genealogy primarily influence the patterns of variance under other conditions. It is the patterns of variance and correlation that determine outcome.

Reasons to favour kin selection over group selection

Kin selection and group selection follow the same partition of variance within and among groups. A group selection analysis tends to emphasize the variance among groups and therefore the effect of selection at the group level. A kin selection analysis tends to emphasize the correlation between members of the same group, measured by the kin selection coefficient. The correlation within groups and the relative amount of variation among groups are simply alternative ways of expressing the partitioning of variances (Frank, 1986).

In more complicated biological problems, it often becomes difficult to express all of the selective forces in terms of relative variances among groups. The problem is that patterns of interaction may differ with respect to different processes, such as mating, competition between certain individuals such as males and competition between other individuals such as females. In that sort of realistic scenario, it is far easier to trace pathways of causation through a series of partial correlations that can be interpreted as an extended form of kin selection analysis (Frank, 1986, 1998). In practice, it is rarely

sensible to express such multiple pathways of causation by expressions of relative amounts of variance among groups, although such expressions may be possible mathematically. For that reason, kin selection often becomes a more natural form of analysis for realistic biological problems, leading to a generalized path analysis framework.

The present article is about the separation between selection and transmission rather than a general approach to pathways of causation. Frank (1997b, 1998) summarized the path analysis approach, although some readers may find those publications a bit technical. I will return to the path analysis methods in a later article in this series.

Short-sighted parasite evolution

Within-group competitiveness often evolves, even though competitiveness reduces the equilibrium fitness of all individuals. The models in the previous sections provided examples. In those models, the favoured value of competitiveness was given in eqn (18) by $z^* = 1 - r$. Competitiveness rises as relatedness between group members, r, declines. The equilibrium fitness from eqn (16) is $w = 1 - z^* = r$. Thus, reduced relatedness in groups increases competitiveness and causes a decline in fitness for all individuals and all groups.

I mentioned one interpretation of this simple model in terms of parasite virulence. Parasites may compete for resources within the host. Greater competitiveness may lead to overexploitation of the host, harming the host and ultimately damaging or destroying the resource on which the parasites depend. In that regard, reduced relatedness of parasites within hosts may lead to enhanced competition and greater virulence, where 'virulence' means the degree of harm the parasites cause the host.

Levin & Bull (1994) emphasized the key role of evolutionary timescale. A long period of within-host evolution, with many rounds of parasite competition and selection, may favour the origin and spread of increasing competitiveness between parasites, leading to greater virulence. That evolution of increasing virulence occurs during the time of an infection within a single host. Such evolutionary increase of virulence can kill the host and, in consequence, kill the parasites themselves. In that regard, the newly evolved virulence is short-sighted, because it provides a local advantage to the parasites in the short run but leads to their extinction in the long run.

If the highly virulent forms that evolve within the host rarely transmit to other hosts, then two distinct timescales exist. On the short timescale within hosts, high virulence repeatedly evolves but does not contribute to the long run evolution of the population. On the long timescale in the population of parasites across hosts, the less virulent forms transmit between hosts better than do the highly virulent forms, causing a moderate to low level of virulence among infective parasites entering a host.

By contrast, if the highly virulent forms that evolve within hosts often transmit to other hosts, then the shorter and longer timescales interact. The short-term evolutionary increase of competitiveness within the host contributes to a transmission bias on the longer timescale. The contribution of the short-term increase in virulence within hosts to the longer timescale depends on the fraction of parasites transmitted between hosts that come from the later population within the host. The next section provides an example.

Demography, timescale and microbial metabolism

In this section, I consider groups that continuously produce transmissible forms. The longer the time for evolution within groups, the greater the transmission bias towards characters favoured within groups. For example, within-group selection often favours greater competitiveness against neighbours. If many generations of selection occur within groups, the greater short-term pressure for competitiveness within groups ultimately increases the competitiveness across all groups.

Microbial metabolism nicely illustrates aspects of timescale (Frank, 2010). Extra energy devoted to resource acquisition speeds metabolic rate and competitive success against neighbours but reduces net efficiency and yield. Thus, the local benefit for rapid resource acquisition trades off against lower yield and reduced competitive success of a group against other groups (Pfeiffer *et al.*, 2001).

Once again, we have a situation in which selection within groups favours more competitive traits, whereas selection between groups favours greater restraint and higher group productivity. The balance between opposing forces ultimately depends on the relative selective bias between groups compared with the transmission bias caused by selection within groups.

An example

Suppose that individual microbial colonies occur in separated patches. Each patch lasts for a while but eventually disappears. During a patch's lifespan, there is a continual flow of resources available to the microbes. The microbes compete for the resources within the patch. Competition occurs by the rate of resource uptake. Individuals that invest more energy in uptake outcompete neighbours for resources, but their net conversion of resource into reproduction is lower because they spend more on uptake rather than productivity. Groups that have highly competitive strains, devoting much energy to competitive increases in resource uptake, have low net productivity.

Colonies continuously send out migrants in proportion to group productivity. Transmission bias occurs when the average competitive trait of migrants differs from the

average trait value among those microbes that founded the colony. The local processes of competition, selection and production of migrants continue until colony extinction. Colony formation and colony extinction set the global birth–death process.

The overall scenario is roughly similar to a host–parasite situation, in which resource patches are like hosts, and parasites send out transmissible progeny continuously from an infected host. Many variations are possible. However, the basic set-up provides a useful expression for the interactions between colony demography and the different timescales of selection bias and transmission bias.

I use the particular assumptions and results of Frank (2010). The interpretation follows the same type of selection–transmission balance of previous sections. However, the earlier models were often designed explicitly to illustrate the partition between selective bias and transmission bias. The value here arises from the more realistic biology, which forces us to parse the components of evolutionary change without the advantage of a toy model designed to give us a simple partition.

No transmission bias

Figure 3 shows the net outcome of selective bias between groups and transmission bias within groups. Each colony forms by a small group of genetically identical cells. When there is no mutation, as shown in the bottom curve, no selection can occur within the colony because there is no genetic variation. Thus, the bottom curve reflects the pure effects of selective bias between groups in the absence of transmission bias. The character, z, is the fraction of energy devoted to resource acquisition relative to the fraction $1 - z$ devoted to reproduction. The character value at which equilibrium occurs is z^*.

To understand the consequences of a pure selective bias between groups, recall from eqn (1) that the total change in a character is

$$\bar{w}\Delta\bar{z} = \Delta S + \Delta\tau,$$

the sum of the change caused by selective bias, ΔS, and transmission bias, $\Delta\tau$. Here, the biases are measured with respect to microbial groups living in isolated patches.
The character value settles to equilibrium when $\bar{w}\Delta\bar{z} = \Delta S + \Delta\tau = 0$. If there is no genetic variation among the initial microbes that start each colony, and no mutation, then there can be no selection within groups and no transmission bias, thus $\Delta\tau = 0$. With no transmission bias, the system comes to equilibrium when $\Delta S = 0$. Put another way, group productivity, which determines the selective bias between groups, ΔS, sets the trade-off between rate and yield. In the lower curve of Fig. 3 with no mutation, the value z^* maximizes yield and leads to $\Delta S = sV = 0$.

In this particular model, one cannot write a simple expression for the balance between rate and yield.

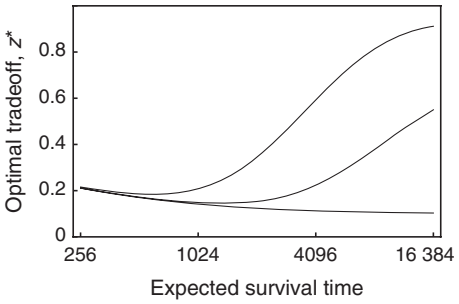

Fig. 3 The trade-off between rate and yield in microbial metabolism. The optimal trade-off, z^*, is the fraction of available resources invested in increasing the rate of acquiring new resources. The remainder of resources, $1 - z^*$, enhances reproduction. The colony survives each time period at rate δ; the expected survival time is $1/\delta$. Each colony begins with a single immigrant or small group of genetically identical immigrants. The microbes use the local resources to reproduce. Mutations occur in the trade-off, z, between rate and yield. The lower curve represents no mutation in the colony. The middle curve has mutation rate, μ, and the upper curve has a higher mutation rate of 10 μ. The colony sends out migrants to colonize new patches. The number of migrants per unit time for each genetic type in a patch is proportional to the number of cells of that genetic type. The details about rate processes are in Frank (2010). Redrawn from fig. 2a of Frank (2010).

Roughly, the idea is that a fraction z of energy is put into increasing the rate of resource acquisition, and a fraction $1 - z$ is put into reproduction or yield. If the factors simply multiplied, then fitness would be $w = z(1 - z)$. The change in fitness with the character z gives the selective coefficient, s. The change in fitness with character value is the derivative of w with respect to z, which gives $s = 1 - 2z = 0$, and so $\Delta S = sV = 0$ implies $z^* = 1/2$.

In the actual model, the length of colony survival affects the balance between rate and yield. Short-lived colonies are favoured to grow quickly (high rate and low efficiency) to use up available resources before extinction, whereas long-lived colonies are favoured to grow slowly and use resources efficiently. Thus, in the lower curve of Fig. 3, longer colony survival causes the optimal balance to shift towards lower rate and higher yield.

Balance between selection and transmission

When mutation generates variation within colonies, the rate–yield trade-off balances selection between colonies and the transmission bias from selection within colonies. The upper two curves in Fig. 3 show the equilibrium balance, z^*. The top curve has a mutation rate ten times greater than the middle curve.

As colony survival increases, the equilibrium moves towards greater investment in resource acquisition. Higher resource acquisition and metabolic rate arise from the inevitable production of mutant neighbours within colonies and the multiple rounds of internal selection

within groups. With very long colony survival times, both upper curves would converge to a high value of z^* at which nearly all resources are devoted to resource acquisition and competition within colonies, with the yield efficiency dropping to a very low level. At that point, transmission bias from selection within groups dominates selection bias between groups.

The equilibrium rate–yield trade-off reflects the fundamental balance between selection and transmission. That balance provides a simple conceptual basis for understanding how natural selection shapes characters. However, in this relatively realistic model, one cannot use the balance of eqn (3) directly to calculate the predicted outcome. Instead, I had to use other mathematical methods to obtain the solution (Frank, 2010). The selection–transmission balance only provides a framing in which to interpret the results.

In the earlier models in this article, it was easy to calculate the ratio of selection to transmission, $\mathscr{R}$. Here, the calculation is difficult, and methods such as the Price equation, kin selection and group selection are of no use in calculating the outcome. After obtaining a solution by other means, one can use those framings to interpret the forces that shaped the outcome. This limitation to post hoc explanations is typical of the grand theories when faced with realistic scenarios. Across the range of different problems presented in this article, the selection versus transmission framing provides the most general conceptual view, following the spirit of the epigraph by Williams.

Conclusions

This article is about the relative contributions of selective bias and transmission bias to overall evolutionary change. For any problem, we first choose a higher level of organization, such as a group, an individual or a cell within a multicellular aggregation. Selective bias arises from differing success among the higher-level entities. Transmission bias arises from changes in character values between higher-level entities and their descendants. Transmission biases may occur by mutation, by random fluctuations and by selection within the group.

The ratio of selective bias to transmission bias provides a simple measure for the relative dominance of the higher to the lower level of organization in overall evolutionary change. When the two levels oppose each other, the relative dominance of one level over the other often sets the level at which functional coherence and individuality emerge.

A key aspect of Maynard Smith & Szathmáry (1995) *The Major Transitions in Evolution* was expressed by Maynard Smith (1988, pp. 229–230):

> One can recognize in the evolution of life several revolutions in the way in which genetic information is organized. In each of these revolutions, there has been a conflict between selection at several levels. The achievement of individuality at the higher level has required that the disruptive effects of selection at the lower level be suppressed.

Maynard Smith's suppression of disruptive effects at the lower level causes selective bias at the higher level to dominate. The quote and the conceptual basis of the major transitions therefore express Williams' notion of the ratio of selective bias to endogenous rate of change (Michod, 1997a; Michod & Nedelcu, 2003).

There is a large philosophical literature on the meaning of *individuality* and of *units of selection* in relation to *levels of selection* (Sober & Wilson, 1994; Okasha, 2006). One can certainly learn from studying that philosophical literature. However, I have found it more instructive to analyse a wide range of interesting biological problems, to discover in practice what is actually needed to understand those problems, and to learn what general concepts link the different problems within a common conceptual basis (cf. Michod, 1997b, 2006). Philosophical induction from numerous evolutionary deductions.

Acknowledgments

My research is supported by National Science Foundation grant EF-0822399, National Institute of General Medical Sciences MIDAS Program grant U01-GM-76499 and a grant from the James S. McDonnell Foundation.

References

Alexander, R.D. & Borgia, G. 1978. Group selection, altruism, and the levels of organization of life. *Annu. Rev. Ecol. Syst.* **9**: 449–474.

Barton, N.H. & Turelli, M. 1987. Adaptive landscapes, genetic distance and the evolution of quantitative characters. *Genet. Res.* **49**: 157–173.

Burt, A. & Trivers, R. 2008. *Genes in Conflict: the Biology of Selfish Genetic Elements*. Belknap Press, Cambridge, MA.

Eigen, M. 1971. Self-organization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **58**: 465–523.

Eigen, M. 1992. *Steps Towards Life: A Perspective on Evolution*. Oxford University Press, Oxford.

Eigen, M. & Schuster, P. 1977. The hypercycle. A principle of natural self-organization. Part A: emergence of the hypercycle. *Naturwissenschaften* **64**: 541–565.

Eigen, M., McCaskill, J. & Schuster, P. 1988. Molecular quasi-species. *J. Phys. Chem.* **92**: 6881–6891.

Eigen, M., McCaskill, J. & Schuster, P. 1989. The molecular quasi-species. *Adv. Chem. Phys.* **75**: 149–263.

Fisher, R.A. 1930. *The Genetical Theory of Natural Selection*. Clarendon, Oxford.

Frank, S.A. 1986. Hierarchical selection theory and sex ratios I. General solutions for structured populations. *Theor. Popul. Biol.* **29**: 312–342.

Frank, S.A. 1987. Demography and sex ratio in social spiders. *Evolution* **41**: 1267–1281.

Frank, S.A. 1994. Kin selection and virulence in the evolution of protocells and parasites. *Proc. R. Soc. Lond. B* **258**: 153–161.
Frank, S.A. 1995a. George Price's contributions to evolutionary genetics. *J. Theor. Biol.* **175**: 373–388.
Frank, S.A. 1995b. Mutual policing and repression of competition in the evolution of cooperative groups. *Nature* **377**: 520–522.
Frank, S.A. 1996a. Host control of symbiont transmission: the separation of symbionts into germ and soma. *Am. Nat.* **148**: 1113–1124.
Frank, S.A. 1996b. Host-symbiont conflict over the mixing of symbiotic lineages. *Proc. R. Soc. Lond. B* **263**: 339–344.
Frank, S.A. 1996c. Models of parasite virulence. *Q. Rev. Biol.* **71**: 37–78.
Frank, S.A. 1997a. Models of symbiosis. *Am. Nat.* **150**: S80–S99.
Frank, S.A. 1997b. The Price equation, Fisher's fundamental theorem, kin selection, and causal analysis. *Evolution* **51**: 1712–1729.
Frank, S.A. 1998. *Foundations of Social Evolution*. Princeton University Press, Princeton, NJ.
Frank, S.A. 2010. The trade-off between rate and yield in the design of microbial metabolism. *J. Evol. Biol.* **23**: 609–613.
Frank, S.A. & Slatkin, M. 1990. The distribution of allelic effects under mutation and selection. *Genet. Res.* **55**: 111–117.
Gardner, A. 2008. The Price equation. *Curr. Biol.* **18**: R198–R202.
Gardner, A. & Grafen, A. 2009. Capturing the superorganism: a formal theory of group adaptation. *J. Evol. Biol.* **22**: 659–671.
Grafen, A. 1984. Natural selection, kin selection and group selection. In: *Behavioural ecology* (J.R. Krebs & N.B. Davies, eds), pp. 62–84. Blackwell Scientific Publications, Oxford.
Grafen, A. 2002. A first formal link between the price equation and an optimization program. *J. Theor. Biol.* **217**: 75–91.
Haldane, J.B.S. 1927. A mathematical theory of natural and artificial selection, Part V: Selection and mutation. *Math. Proc. Cambridge Philos. Soc.* **23**: 838–844.
Hamilton, W.D. 1970. Selfish and spiteful behaviour in an evolutionary model. *Nature.* **228**: 1218–1220.
Hamilton, W.D. 1975. Innate social aptitudes of man: an approach from evolutionary genetics. In: *Biosocial Anthropology* (R. Fox, ed.), pp. 133–155. Wiley, New York.
Hamilton, W.D. 1979. Wingless and fighting males in fig wasps and other insects. In: *Reproductive Competition and Sexual Selection in Insects* (M.S. Blum & N.A. Blum, eds), pp. 167–220. Academic Press, New York.
Heisler, I.L. & Damuth, J. 1987. A method for analyzing selection in hierarchically structured populations. *Am. Nat.* **130**: 582–602.
Kimura, M. 1965. A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proc. Natl Acad. Sci. USA* **54**: 731–736.
Lande, R. 1975. The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genet. Res.* **26**: 221–235.
Levin, B.R. & Bull, J.J. 1994. Short-sighted evolution and the virulence of pathogenic microorganisms. *Trends. Microbiol.* **2**: 76–81.
Lewontin, R.C. 1970. The units of selection. *Annu. Rev. Ecol. Syst.* **1**: 1–18.
Lynch, M., Bürger, R., Butcher, D. & Gabriel, W. 1993. The mutational meltdown in asexual populations. *J. Hered.* **84**: 339–344.
Maynard Smith, J. 1976. Group selection. *Q. Rev. Biol.* **51**: 277–283.
Maynard Smith, J. 1979. Hypercycles and the origin of life. *Nature.* **280**: 445–446.
Maynard Smith, J. 1988. Evolutionary progress and levels of selection. In: *Evolutionary Progress* (M.H. Nitecki, ed.), pp. 219–230. University of Chicago Press, Chicago.
Maynard Smith, J. & Szathmáry, E. 1995. *The Major Transitions in Evolution*. Freeman, San Francisco.
Michod, R.E. 1997a. Cooperation and conflict in the evolution of individuality. I. Multilevel selection of the organism. *Am. Nat.* **149**: 607–645.
Michod, R.E. 1997b. Evolution of the individual. *Am. Nat.* **150**: S5–S21.
Michod, R.E. 2006. The group covariance effect and fitness trade-offs during evolutionary transitions in individuality. *Proc. Natl Acad. Sci. USA* **103**: 9113–9117.
Michod, R.E. & Herron, M.D. 2006. Cooperation and conflict during evolutionary transitions in individuality. *J. Evol. Biol.* **19**: 1406–1409.
Michod, R.E. & Nedelcu, A.M. 2003. On the reorganization of fitness during evolutionary transitions in individuality. *Integr. Comp. Biol.* **43**: 64–73.
Nachman, M. & Crowell, S. 2000. Estimate of the mutation rate per nucleotide in humans. *Genetics.* **156**: 297–304.
Okasha, S. 2006. *Evolution and the Levels of Selection*. Oxford University Press, New York.
Pfeiffer, T., Schuster, S. & Bonhoeffer, S. 2001. Cooperation and competition in the evolution of ATP-producing pathways. *Science* **292**: 504–507.
Price, G.R. 1970. Selection and covariance. *Nature* **227**: 520–521.
Price, G.R. 1972. Extension of covariance selection mathematics. *Ann. Hum. Genet.* **35**: 485–490.
Queller, D.C. 1994. Genetic relatedness in viscous populations. *Evol. Ecol.* **8**: 70–73.
Rice, S.H. 2004. *Evolutionary Theory: Mathematical and Conceptual Foundations*. Sinauer Associates, Sunderland, MA.
Sober, E. & Wilson, D.S. 1994. A critical review of philosophical work on the units of selection problem. *Philos. Sci.* **61**: 534–555.
Stearns, S.C. 1986. Natural selection and fitness, adaptation and constraint. In: *Patterns and Processes in the History of Life* (D. Jablonski & D.M. Raup, eds), pp. 23–44. Springer-Verlag, Berlin.
Stebbins, G.L. 1950. *Variation and Evolution in Plants*. Columbia University Press, New York.
Szathmáry, E. & Demeter, L. 1987. Group selection of early replicators and the origin of life. *J. Theor. Biol.* **128**: 463–486.
Taylor, P.D. 1992. Altruism in viscous populations – an inclusive fitness approach. *Evol. Ecol.* **6**: 352–356.
Turelli, M. 1984. Heritable genetic variation via mutation–selection balance: Lerch's zeta meets the abdominal bristle. *Theor. Popul. Biol.* **25**: 138–193.
Van Valen, L.M. 1975. Group selection, sex, and fossils. *Evolution* **29**: 87–94.
Van Valen, L.M. 1988. Species, sets, and the derivative nature of philosophy. *Biol. Philos.* **3**: 49–66.
Wade, M.J. 1985. Soft selection, hard selection, kin selection, and group selection. *Am. Nat.* **125**: 61–73.
West, S.A., Griffin, A.S. & Gardner, A. 2007. Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *J. Evol. Biol.* **20**: 415–432.

Wilke, C. 2005. Quasispecies theory in the context of population genetics. *BMC Evol. Biol.* **5**: 44.
Williams, G.C. 1966. *Adaptation and Natural Selection*. Princeton University Press, Princeton, NJ.
Williams, G.C. 1992. *Natural Selection: Domains, Levels, and Challenges*. Oxford University Press, Oxford.
Wilson, D.S. & Colwell, R.K. 1981. The evolution of sex ratio in structured demes. *Evolution* **35**: 882–897.
Wilson, D.S. & Sober, E. 1989. Reviving the superorganism. *J. Theor. Biol.* **136**: 337–356.
Wilson, D.S., Pollock, G.B. & Dugatkin, L.A. 1992. Can altruism evolve in purely viscous populations? *Evol. Ecol.* **6**: 331–341.

Received 13 October 2011; accepted 1 November 2011

JOURNAL OF Evolutionary Biology

doi: 10.1111/j.1420-9101.2012.02498.x

REVIEW

Natural selection. IV. The Price equation*

S. A. FRANK

Department of Ecology and Evolutionary Biology, University of California, Irvine, CA, USA

Keywords:

fundamental theorem;
invariance expressions;
mathematical models;
population genetics.

Abstract

The Price equation partitions total evolutionary change into two components. The first component provides an abstract expression of natural selection. The second component subsumes all other evolutionary processes, including changes during transmission. The natural selection component is often used in applications. Those applications attract widespread interest for their simplicity of expression and ease of interpretation. Those same applications attract widespread criticism by dropping the second component of evolutionary change and by leaving unspecified the detailed assumptions needed for a complete study of dynamics. Controversies over approximation and dynamics have nothing to do with the Price equation itself, which is simply a mathematical equivalence relation for total evolutionary change expressed in an alternative form. Disagreements about approach have to do with the tension between the relative valuation of abstract versus concrete analyses. The Price equation's greatest value has been on the abstract side, particularly the invariance relations that illuminate the understanding of natural selection. Those abstract insights lay the foundation for applications in terms of kin selection, information theory interpretations of natural selection and partitions of causes by path analysis. I discuss recent critiques of the Price equation by Nowak and van Veelen.

> The heart and soul of much mathematics consists of the fact that the 'same' object can be presented to us in different ways. Even if we are faced with the simple-seeming task of 'giving' a large number, there is no way of doing this without also, at the same time, 'giving' a hefty amount of extra structure that comes as a result of the way we pin down—or the way we present—our large number. If we write our number as 1729, we are, sotto voce, ordering a preferred way of 'computing it' (add one thousand to seven hundreds to two tens to nine). If we present it as $1 + 12^3$, we are recommending another mode of computation, and if we pin it down—as Ramanujuan did—as the first number expressible as a sum of two cubes in two different ways, we are being less specific about how to compute our number, but have underscored a characterizing property of it within a subtle diophantine arena.…
>
> This issue has been with us, of course, forever: the general question of *abstraction*, as separating what we want from what we are presented with. It is neatly packaged in the Greek verb *aphairein*, as interpreted by Aristotle in the later books of the *Metaphysics* to mean simply *separation*: if it is *whiteness* we want to think about, we must somehow separate it from *white horse*, *white house*, *white hose* and all the other white things that it invariably *must* come along with, in order for us to experience it at all
>
> (Mazur, 2008, pp. 222–223).

> Somewhere … between the specific that has no meaning and the general that has no content there must be, for each purpose and at each level of abstraction, an optimum degree of generality
>
> (Boulding, 1956, pp. 197–198).

Introduction

Evolutionary theory analyses the change in phenotype over time. We may interpret *phenotype* broadly to include organismal characters, variances of characters, correlations between characters, gene frequency, DNA sequence – essentially anything we can measure.

How does a phenotype influence its own change in frequency or the change in the frequencies of correlated phenotypes? Can we separate that phenotypic influence

Correspondence: S. A. Frank, Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697–2525, USA.
Tel.: +1 949 824 2244; fax: +1 949 824 2181; e-mail: safrank@uci.edu

*Part of the Topics in Natural Selection series. See Box 1.

Box 1: Topics in the theory of natural selection

This article is part of a series on natural selection. Although the theory of natural selection is simple, it remains endlessly contentious and difficult to apply. My goal is to make more accessible the concepts that are so important, yet either mostly unknown or widely misunderstood. I write in a nontechnical style, showing the key equations and results rather than providing full derivations or discussions of mathematical problems. Boxes list technical issues and brief summaries of the literature.

from other evolutionary forces that also cause change? The association of a phenotype with change in frequency, separated from other forces that change phenotype, is one abstract way to describe natural selection. The Price equation is that kind of abstract separation.

Do we really need such abstraction, which may seem rather distant and vague? Instead of wasting time on such things as the abstract essence of natural selection, why not get down to business and analyse real problems? For example, we may wish to know how the evolutionary forces of mutation and selection interact to determine biological pattern. We could make a model with genes that have phenotypic effects, selection that acts on those phenotypes to change gene frequency, and mutation that changes one gene into another. We could do some calculations, make some predictions about, for example, the frequency of deleterious mutations that cause disease, and compare those predictions to observations. All clear and concrete, without need of any discussion of the essence of things.

However, we may ask the following. Is there some reorientation for the expression of natural selection that may provide subtle perspective, from which we can understand our subject more deeply and analyse our problems with greater ease and greater insight? My answer is, as I have mentioned, that the Price equation provides that sort of reorientation. To argue the point, I will have to keep at the distinction between the concrete and the abstract, and the relative roles of those two endpoints in mature theoretical understanding.

Several decades have passed since Price's (1970, 1972a) original articles. During that span, published claims, counter-claims and misunderstandings have accumulated to the point that it seems worthwhile to revisit the subject. On the one hand, the Price equation has been applied to numerous practical problems and has also been elevated by some to almost mythical status, as if it were the ultimate path to enlightenment for those devoted to evolutionary study (Box 2).

On the other hand, the opposition has been gaining adherents who boast the sort of disparaging anecdotes and slogans that accompany battle. In a recent book, Nowak & Highfield (2011) counter

Box 2: Price equation literature

A large literature introduces and reviews the Price equation. I list some key references that can be used to get started (Hamilton, 1975; Frank, 1995, 1997; Grafen, 2002; Page & Nowak, 2002; Andersen, 2004; Rice, 2004; Okasha, 2006; Gardner, 2008).

Diverse applications have been developed with the Price equation. I list a few examples (Hamilton, 1970; Wade, 1985; Frank & Slatkin, 1990; Queller, 1992a,b; Michod, 1997a,b; Frank, 1998; Day & Gandon, 2006; Fox, 2006; Grafen, 2007; Alizon, 2009).

Quantitative genetics theory often derives from the covariance expression given by Robertson (1966), which is a form of the covariance term of the Price equation. The basic theory can be found in textbooks (Falconer & Mackay, 1996; Charlesworth & Charlesworth, 2010). Much of the modern work can be traced through the widely cited article by Lande & Arnold (1983).

Harman (2010) provides an interesting overview of Price's life and evokes an Olympian sense of the power and magic of the Price equation. See the study of Schwartz (2000) for an alternative biographical sketch.

> The Price equation did not, however, prove as useful as [Price and Hamilton] had hoped. It turned out to be the mathematical equivalent of a tautology. … If the Price equation is used instead of an actual model, then the arguments hang in the air like a tantalizing mirage. The meaning will always lie just out of the reach of the inquisitive biologist. This mirage can be seductive and misleading. The Price equation can fool people into believing that they have built a mathematical model of whatever system they are studying. But this is often not the case. Although answers do indeed seem to pop out of the equation, like rabbits from a magician's hat, nothing is achieved in reality.

Nowak & Highfield (2011) approvingly quote van Veelen *et al.* (2012) with regard to calling the Price equation a *mathematical tautology*. van Veelen *et al.* (2012) emphasize the point by saying that the Price equation is like soccer/football star Johan Cruyff's quip about the secret of success: 'You always have to make sure that you score one goal more than your opponent'. The statement is always true, but provides no insight. Nowak & Highfield (2011) and van Veelen *et al.* (2012) believe their arguments demonstrate that the Price equation is true in the same trivial sense, and they call that trivial type of truth a *mathematical tautology*. Interestingly, magazines, online articles and the scientific literature have for several years been using the phrase *mathematical tautology* for the Price equation, although Nowak & Highfield (2011) and van Veelen *et al.* (2012) do not provide citations to previous literature.

As far as I know, the first description of the Price equation as a mathematical tautology was in the study of Frank (1995). I used the phrase in the sense of the epigraph from Mazur, a formal equivalence between

different expressions of the same object. Mathematics and much of statistics are about formal equivalences between different expressions of the same object. For example, the Laplace transform changes a mathematical expression into an alternative form with the same information, and analysis of variance decomposes the total variance into a sum of component variances. For any mathematical or statistical equivalence, value depends on enhanced analytical power that eases further derivations and calculations and on the ways in which previously hidden relations are revealed.

In the light of the contradictory points of view, the main goal of this article is to sort out exactly what the Price equation is, how we should think about it, and its value and limitations in reasoning about evolution. Subsequent articles will show the Price equation in action, applied to kin selection, causal analysis in evolutionary models and an information perspective of natural selection and Fisher's fundamental theorem.

Overview

The first section derives the Price equation in its full and most abstract form. That derivation allows us to evaluate the logical status of the equation in relation to various claims of fundamental flaw. The equation survives scrutiny. It is a mathematical relation that expresses the total amount of evolutionary change in an alternative and mathematically equivalent way. That equivalence provides insight into aspects of natural selection and also provides a guide that, in particular applications, often leads to good approaches for analysis.

The second section contrasts two perspectives of evolutionary analysis. In standard models of evolutionary change, one begins with the initial population state and the rules of change. The rules of change include the fitness of each phenotype and the change in phenotype between ancestor and descendant. Given the initial state and rules of change, one deduces the state of the changed population. Alternatively, one may have data on the initial population state, the changed population state and the ancestor–descendant relations that map entities from one population to the other. Those data may be reduced to the evolutionary distance between two populations, providing inductive information about the underlying rules of change. Natural populations have no intrinsic notion of fitness or rules of change. Instead, they inductively accumulate information. The Price equation includes both the standard deductive model of evolutionary change and the inductive model by which information accumulates in relation to the evolutionary distance between populations.

The third and fourth sections discuss the Price equation's abstract properties of invariance and recursion. The invariance properties include the information theory interpretation of natural selection. Recursion provides the basis for analysing group selection and other models of multilevel selection.

The fifth section relates the Price equation to various expressions that have been used throughout the history of evolutionary theory to analyse natural selection. The most common form describes natural selection by the covariance between phenotype and fitness or by the covariance between genetic breeding value and fitness. The covariance expression is one part of the Price equation that, when used alone, describes the natural selection component of total evolutionary change. The essence of those covariance forms arose in the early studies of population and quantitative genetics, had been used extensively during much of the modern history of animal breeding, and began to receive more mathematical development in the 1960s and 1970s. Recent critiques of the Price equation focus on the same covariance expression that has been widely used throughout the history of population and quantitative genetics to analyse natural selection and to approximate total evolutionary change.

The sixth section returns to the full abstract form of the equation. I compare a few variant expressions that have been promoted as improvements on the original Price equation. Variant forms are indeed helpful with regard to particular abstract problems or particular applications. However, most variants are simply minor rearrangements of the mathematical equivalence for total evolutionary change given by the original Price equation. The recent extension by Kerr & Godfrey-Smith (2009) does provide a slightly more general formulation by expanding the fundamental set mapping that defines Price's approach. The set mapping basis for the Price equation deserves more careful study and further mathematical work.

The seventh section analyses various flaws that have been ascribed to the Price equation. For example, the Price equation in its most abstract form does not contain enough information to follow evolutionary dynamics through multiple rounds of natural selection. By contrast, classical dynamic models of population genetics are sufficient to follow change through time. Much has been made of this distinction with regard to dynamic sufficiency. The distinction arises from the fact that classical dynamics in population genetics makes more initial assumptions than the abstract Price equation. It must be true that all mathematical equivalences for total evolutionary change have the same dynamic status given the same initial assumptions. Each additional well-chosen assumption typically enhances the specificity and reduces the scope and generality of the analysis. The epigraph from Boulding emphasizes that the degree of specificity versus generality is an explicit choice of the analyst with respect to initial assumptions.

The Discussion considers the value and limitations of the Price equation in relation to recent criticisms by Nowak and van Veelen. The critics confuse the distinct roles of general abstract theory and concrete dynamical models for particular cases. The enduring power of the Price equation arises from the discovery of essential invariances in natural selection. For example, kin selection theory expresses biological problems in terms of relatedness coefficients. Relatedness measures the association between social partners. The proper measure of relatedness identifies distinct biological scenarios with the same (invariant) evolutionary outcome. Invariance relations provide the deepest insights of scientific thought.

The Price equation

> The mathematics given here applies not only to genetical selection but to selection in general. It is intended mainly for use in deriving general relations and constructing theories, and to clarify understanding of selection phenomena, rather than for numerical calculation (Price, 1972a, p. 485).

I have emphasized that the Price equation is a mathematical equivalence. The equation focuses on separation of total evolutionary change into a part attributed to selection and a remainder term. That separation provides an abstraction of the nature of selection. As Price wrote sometime around 1970 but published posthumously in Price (1995), 'Despite the pervading importance of selection in science and life, there has been no abstraction and generalization from genetical selection to obtain a general selection theory and general selection mathematics'.

It is useful first to consider the Price equation in this most abstract form. I follow my earlier derivations (Frank, 1995, 1997, 1998, 2009), which differ little from the derivation given by Price (1972a) when interpreted in the light of the study of Price (1995).

The abstract expression can best be thought of in terms of mapping items between two sets (Frank, 1995; Price, 1995). In biology, we usually think of an ancestral population at some time and a descendant population at a later time. Although there is no need to have an ancestor–descendant relation, I will for convenience refer to the two sets as ancestor and descendant. What does matter is the relations between the two sets, as follows.

Definitions

The full abstract power of the Price equation requires adhering strictly to particular definitions. The definitions arise from the general expression of the relations between two sets.

Let q_i be the frequency of the ith type in the ancestral population. The index i may be used as a label for any sort of property of things in the set, such as allele, genotype, phenotype, group of individuals and so on. Let q_i' be the frequencies in the descendant population, defined as the fraction of the descendant population that is derived from members of the ancestral population that have the label i. Thus, if $i = 2$ specifies a particular phenotype, then q_2' is not the frequency of the phenotype $i = 2$ among the descendants. Rather, it is the fraction of the descendants derived from entities with the phenotype $i = 2$ in the ancestors. One can have partial assignments, such that a descendant entity derives from more than one ancestor, in which case each ancestor gets a fractional assignment of the descendant. The key is that the i indexing is always with respect to the properties of the ancestors, and descendant frequencies have to do with the fraction of descendants derived from particular ancestors.

Given this particular mapping between sets, we can specify a particular definition for fitness. Let $q_i' = q_i(w_i/\bar{w})$, where w_i is the fitness of the ith type and $\bar{w} = \sum q_i w_i$ is average fitness. Here, $w_i/\bar{w}$ is proportional to the fraction of the descendant population that derives from type i entities in the ancestors.

Usually, we are interested in how some measurement changes or evolves between sets or over time. Let the measurement for each i be z_i. The value z may be the frequency of a gene, the squared deviation of some phenotypic value in relation to the mean, the value obtained by multiplying measurements of two different phenotypes of the same entity and so on. In other words, z_i can be a measurement of any property of an entity with label, i. The average property value is $\bar{z} = \sum q_i z_i$, where this is a population average.

The value z_i' has a peculiar definition that parallels the definition for q_i'. In particular, z_i' is the average measurement of the property associated with z among the descendants derived from ancestors with index i. The population average among descendants is $\bar{z}' = \sum q_i' z_i'$.

The Price equation expresses the total change in the average property value, $\Delta\bar{z} = \bar{z}' - \bar{z}$, in terms of these special definitions of set relations. This way of expressing total evolutionary change and the part of total change that can be separated out as selection is very different from the usual ways of thinking about populations and evolutionary change. The derivation itself is very easy, but grasping the meaning and becoming adept at using the equation is not so easy.

I will present the derivation in two stages. The first stage makes the separation into a part ascribed to selection and a part ascribed to property change that covers everything beyond selection. The second stage retains this separation, changing the notation into standard statistical expressions that provide the form of the Price equation commonly found in the literature. I follow with some examples to illustrate how particular set relations are separated into selection and property change components. The next section considers two distinct interpretations of the Price equation in relation to dynamics.

Derivation: separation into selection and property value change

We use $\Delta q_i = q_i' - q_i$ for frequency change associated with selection and $\Delta z_i = z_i' - z_i$ for property value change. Both expressions for change depend on the special set relation definitions given above.

We are after an alternative expression for total change, $\Delta\bar{z}$. Thus,

$$\begin{aligned}\Delta\bar{z} &= \bar{z}' - \bar{z} \\ &= \sum q_i' z_i' - \sum q_i z_i \\ &= \sum q_i'(z_i' - z_i) + \sum q_i' z_i - \sum q_i z_i \\ &= \sum q_i'(\Delta z_i) + \sum (\Delta q_i) z_i.\end{aligned}$$

Switching the order of the terms on the right side of the last line yields

$$\Delta\bar{z} = \sum (\Delta q_i) z_i + \sum q_i'(\Delta z_i), \tag{1}$$

a form emphasized by Frank (1997, eqn 1). The first term separates the part of total change caused by changes in frequency. We call this the part caused by selection, because this is the part that arises directly from differential contribution by ancestors to the descendant population (Price, 1995). Because the set mappings define all of the direct attributions of success for each i with respect to the associated properties z_i, it is reasonable to separate out this direct component as the abstraction of selection. It is of course possible to define other separations. I discuss one particular alternative later. However, it is hard to think of other separations that would describe selection in a better way at the most abstract and general level of the mappings between two sets. This first term has also been called the partial evolutionary change caused by natural selection (eqn 7).

The second term describes the part of total change caused by changes in property values. Recall that $\Delta z_i = z_i' - z_i$ and that z_i' is the property value among entities that descend from i. Many different processes may cause descendant property values to differ from ancestral values. In fact, the assignment of a descendant to an ancestor can be entirely arbitrary, so that there is no reason to assume that descendants should be like ancestors. Usually, we will work with systems in which descendants do resemble ancestors, but the degree of such associations can be arranged arbitrarily. This term for change in property value encompasses everything beyond selection. The idea is that selection affects the relative contribution of ancestors and thus the changes in frequencies of representation, but what actually gets represented among the descendants will be subject to a variety of processes that may alter the value expressed by descendants.

The equation is exact and must apply to every evolutionary system that can be expressed as two sets with certain ancestor–descendant or mapping relations. It is in that sense that I first used the phrase *mathematical tautology* (Frank, 1995). The nature of separation and abstraction is well described by the epigraph from Mazur at the start of this article.

Derivation: statistical notation

Price (1972a) used statistical notation to write eqn 1. For the first term, by following prior definitions, we have

$$\begin{aligned}\Delta q_i &= q_i' - q_i \\ &= q_i \frac{w_i}{\bar{w}} - q_i \\ &= q_i\left(\frac{w_i}{\bar{w}} - 1\right),\end{aligned}$$

so that

$$\sum (\Delta q_i) z_i = \sum q_i \left(\frac{w_i}{\bar{w}} - 1\right) z_i = \mathrm{Cov}(w, z)/\bar{w},$$

using the standard definition for population covariance. For the second term, we have

$$\sum q_i'(\Delta z_i) = \sum q_i \frac{w_i}{\bar{w}} (\Delta z_i) = \mathrm{E}(w\Delta z)/\bar{w},$$

where E means expectation, or average over the full population. Putting these statistical forms into eqn 1 and moving $\bar{w}$ to the left side for notational convenience yields a commonly published form of the Price equation

$$\bar{w}\Delta\bar{z} = \mathrm{Cov}(w, z) + \mathrm{E}(w\Delta z). \tag{2}$$

Frank (1995) and Price (1995) present examples of set mappings expressed in relation to the Price equation.

Dynamics: inductive and deductive perspectives

The Price equation describes evolutionary change between two populations. Three factors express one iteration of dynamical change: initial state, rules of change and next state. In the Price equation, the phenotypes, z_i, and their frequencies, q_i, describe the initial population state. Fitnesses, w_i, and property changes, Δz_i, set the rules of change. Derived phenotypes, z_i', and their frequencies, q_i', express the next population state.

Models of evolutionary change essentially always analyse forward or deductive dynamics. In that case, one starts with initial conditions and rules of change and calculates the next state. Most applications of the Price equation use this traditional deductive analysis. Such applications lead to predictions of evolutionary outcome given assumptions about evolutionary process, expressed by the fitness parameters and property changes.

Alternatively, one can take the state of the initial population and the state of the changed population as given. If one also has the mappings between initial and changed populations that connect each entity, i, in the initial population to entities in the changed population, then one can calculate (induce) the underlying rules of

change. At first glance, this inductive view of dynamics may seem rather odd and not particularly useful. Why start with knowledge of the evolutionary sequence of population states and ancestor–descendant relations as given, and inductively calculate fitnesses and property changes? The inductive view takes the fitnesses, w_i, to be derived from the data rather than an intrinsic property of each type.

The Price equation itself does not distinguish between the deductive and inductive interpretations. One can specify initial state and rules of change and then deduce outcome. Or one can specify initial state and outcome along with ancestor–descendant mappings and then induce the underlying rules of change. It is useful to understand the Price equation in its full mathematical generality and to understand that any specific interpretation arises from additional assumptions that one brings to a particular problem. Much of the abstract power of the Price equation comes from understanding that, by itself, the equation is a minimal description of change between populations.

The deductive interpretation of the Price equation is clear. What value derives from the inductive perspective? In observational studies of evolutionary change, we only have data on population states. From those data, we use the inductive perspective to make inferences about the underlying rules of change. Note that inductive estimates for evolutionary process derive from the amount of change, or distance, between ancestor and descendant populations. The Price equation includes that inductive, or retrospective, view by expressing the distance between populations in terms of $\Delta\bar{z}$. I develop that distance interpretation in the following sections.

Perhaps more importantly, natural selection itself is inherently an inductive process by which information accumulates in populations. Nature does not intrinsically 'know' of fitness parameters. Instead, frequency changes and the mappings between ancestor and descendant are inherent in a population's response to the environment, leading to a sequence of population states, each separated by an evolutionary distance. That evolutionary distance provides information that populations accumulate inductively about the fitnesses of each phenotype (Frank, 2009). The Price equation includes both the deductive and inductive perspectives. We may choose to interpret the equation in either way depending on our goals of analysis.

Abstract properties: invariance

The Price equation describes selection by the term $\sum(\Delta q_i)z_i = \mathrm{Cov}(w,z)/\bar{w}$. Any instance of evolutionary change that has the same value for this sum has the same amount of total selection. Put another way, for any particular value for total selection, there is an infinite number of different combinations of frequency changes and character measurements that will add up to the same total value for selection. All of those different combinations lead to the same value with respect to the amount of selection. We may say that all of those different combinations are *invariant* with respect to the total quantity of selection. The deepest insights of science come from understanding what does not matter, so that one can also say exactly what does matter – what is invariant (Feynman, 1967; Weyl, 1983).

The invariance of selection with respect to transformations of the fitnesses, w, and the phenotypes, z, that have the same $\mathrm{Cov}(w,z)$ means that, to evaluate selection, it is sufficient to analyse this covariance. At first glance, it may seem contradictory that the covariance, commonly thought of as a linear measure of association, can be a complete description for selection, including nonlinear processes. Let us step through this issue, first looking at why the covariance is a sufficient expression of selection and then at the limitations of this covariance expression in evolutionary analysis.

Covariance as a measure of distance: definitions

Much of the confusion with respect to covariance and variance terms in selection equations arises from thinking only of the traditional statistical usage. In statistics, covariance typically measures the linear association between pairs of observations, and variance is a measure of the squared spread of observations. Alternatively, covariances and variances provide measures of distance, which ultimately can be understood as measures of information (Frank, 2009). This section introduces the notation for the geometric interpretation of distance. The next section gives the main geometric result, and the following section presents some examples.

The identity $\sum(\Delta q_i)z_i = \mathrm{Cov}(w,z)/\bar{w}$ provides the key insight. It helps to write this identity in an alternative form. Note from the prior definition $q_i' = q_i w_i/\bar{w}$ that

$$\Delta q_i = q_i' - q_i = q_i(w_i/\bar{w} - 1) = q_i a_i, \tag{3}$$

where $a_i = w_i/\bar{w} - 1$ is Fisher's average excess in fitness, a commonly used expression in population and quantitative genetics (Fisher, 1930, 1941; Crow & Kimura, 1970). A value of zero means that an entity has average fitness, and therefore, fitness effects and selection do not change the frequency of that entity. Using the average excess in fitness, we can write the invariant expression for selection as

$$\sum(\Delta q_i)z_i = \sum q_i a_i z_i = \mathrm{Cov}(w,z)/\bar{w}. \tag{4}$$

We can think of the state of the population as the listing of character states, z_i. Thus, we write the population state as $\mathbf{z} = (z_1, z_2, \ldots)$. The subscripts run over every different entity in the population, so the vector $\mathbf{z}$ is a complete description of the entire population. Similarly, for the frequency fluctuations, $\Delta q_i = q_i a_i$, we can write the listing of all fluctuations as a vector, $\Delta\mathbf{q} = (\Delta q_1, \Delta q_2, \ldots)$.

It is often convenient to use the dot product notation

$$\Delta\mathbf{q} \cdot \mathbf{z} = \sum(\Delta q_i) z_i = \mathrm{Cov}(w, z)/\bar{w}$$

in which the dot specifies the sum obtained by multiplying each pair of items from two vectors. Before turning to some geometric examples in the following section, we need a definition for the length of a vector. Traditionally, one uses the definition

$$\|\mathbf{z}\| = \sqrt{\sum z_i^2},$$

in which the length is the square root of the sum of squares, which is the standard measure of length in Euclidean geometry.

Covariance as a measure of distance: examples

A simple identity relates a dot product to a measure of distance and to covariance selection

$$\Delta\mathbf{q} \cdot \mathbf{z} = \|\Delta\mathbf{q}\| \, \|\mathbf{z}\| \cos\phi = \mathrm{Cov}(w, z)/\bar{w}, \qquad (5)$$

where ϕ is the angle between the vectors $\Delta\mathbf{q}$ and $\mathbf{z}$ (Fig. 1). If we standardize the character vector $\mathbf{z}^* = \mathbf{z}/\|\mathbf{z}\|$, then the standardized vector has a length of one, $\|\mathbf{z}^*\| = 1$, which simplifies the dot product expression of selection to

$$\Delta\mathbf{q} \cdot \mathbf{z}^* = \|\Delta\mathbf{q}\| \cos\phi,$$

providing the geometric representation illustrated in Fig. 1.

The covariance can be expressed as the product of a regression coefficient and a variance term

$$\mathrm{Cov}(w, z)/\bar{w} = \beta_{zw}\mathrm{Var}(w)/\bar{w} = \beta_{wz}\mathrm{Var}(z)/\bar{w}, \qquad (6)$$

where the notation β_{xy} describes the regression coefficient of x on y (Price, 1970). This identity shows that the expression of selection in terms of a regression coefficient and a variance term is equivalent to the geometric expression of selection in terms of distance.

I emphasize these identities for two reasons. First, as Mazur stated in the epigraph, 'The heart and soul of much mathematics consists of the fact that the ''same'' object can be presented to us in different ways'. If an object is important, such as natural selection surely is, then it pays to study that object from different perspectives to gain deeper insight.

Second, the appearance of statistical functions, such as the covariance and variance, in selection equations sometimes leads to mistaken conclusions. In the selection equations, it is better to think of the covariance and variance terms arising because they are identities with geometric or other interpretations of selection, rather than thinking of those terms as summary statistics of probability distributions. The problem with thinking of those terms as statistics of probability distributions is that the variance and covariance are not in general sufficient descriptions for probability distributions. That lack of sufficiency for probability may lead one to conclude that those terms are not sufficient for a general expression of selection. However, those covariance and variance terms are sufficient. That sufficiency can be understood by thinking of those terms as identities for distance or measures of information (Frank, 2009).

It is true that in certain particular applications of quantitative genetics or stochastic sampling processes,

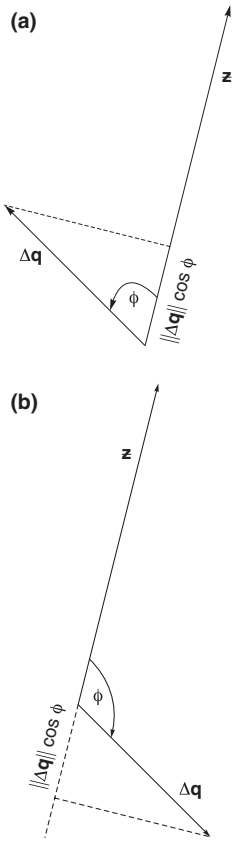

Fig. 1 Geometric expression of selection. The plots show the equivalence of the dot product, the geometric expression and the covariance, as given in eqn 5. For both plots, $\mathbf{z} = (1,4)$ and $\mathbf{z}^* = \mathbf{z}/\|\mathbf{z}\| = (0.24, 0.97)$. The dashed line shows the perpendicular between the pattern of frequency changes derived from fitnesses, $\Delta\mathbf{q}$, and the phenotypic pattern, $\mathbf{z}^*$. The vertex of the two vectors is at the origin (0,0). The distance from the origin to the intersection of the perpendicular along $\mathbf{z}^*$ is the total amount of selection, $\|\Delta\mathbf{q}\| \cos\phi$. (a) The vector of frequency changes that summarize fitness is $\Delta\mathbf{q} = (-0.4, 0.4)$. The angle between the vector of frequency changes and the phenotypes is $\phi = \arccos[(\Delta\mathbf{q} \cdot \mathbf{z}^*)/\|\Delta\mathbf{q}\|]$ which, in this example, is 1.03 radians or 59°. In this case, the total selection is $\|\Delta\mathbf{q}\| \cos\phi = 0.29$. (b) In this plot, $\Delta\mathbf{q} = (0.4, -0.4)$, yielding an angle ϕ of 121°. The perpendicular intersects the negative projection of the phenotype vector, shown as a dashed line, associated with the negative change by selection of $\|\Delta\mathbf{q}\| \cos\phi = -0.29$.

one does interpret the variances and covariances as summary statistics of probability distributions, usually the normal or Gaussian distribution. However, it is important to distinguish those special applications from the general selection equations.

Invariance and information

For the general selection expression in eqn 5, any transformations that do not affect the net values are invariant with respect to selection. For example, transformations of the fitnesses and associated frequency changes, $\Delta\mathbf{q}$, are invariant if they leave unchanged the distance expressed by $\Delta\mathbf{q}\cdot\mathbf{z} = \mathrm{Cov}(w,z)/\bar{w}$. Similarly, changes in the pattern of phenotypes are invariant to the extent that they leave $\Delta\mathbf{q}\cdot\mathbf{z}$ unchanged. These invariance properties of selection, measured as distance, may not appear very interesting at first glance. They seem to be saying that the outcome is the outcome. However, the history of science suggests that studying the invariant properties of key expressions can lead to insight.

Few authors have developed an interest in the invariant qualities of selection. Fisher (1930) initiated discussion with his fundamental theorem of natural selection, a special case of eqn 5 (Frank, 1997). Although many authors commented on the fundamental theorem, most articles did not analyse the theorem with respect to its essential mathematical insights about selection. Ewens (1992) reviewed the few attempts to understand the mathematical basis of the theorem and its invariant quantities. Frank (2009) tied the theorem to Fisher information (Frieden *et al.*, 2001; Frieden, 2004), hinting at an information theory interpretation that arises from the fundamental selection equation of eqn 5.

In spite of the importance of selection in many fields of science, the potential interpretation of eqn 5 with respect to invariants of information theory has hardly been developed. I briefly outline the potential connections here (Frank, 2009). I develop this information perspective of selection in a later article, along with Fisher's fundamental theorem.

To start, define the partial change in phenotype caused by natural selection as

$$\Delta_s\bar{z} = \Delta\mathbf{q}\cdot\mathbf{z} = \mathrm{Cov}(w,z)/\bar{w}. \tag{7}$$

The concept of a partial change caused by natural selection arises from Fisher's fundamental theorem (Fisher, 1930; Price, 1972b; Ewens, 1989; Frank & Slatkin, 1992). With this definition, we can use eqns 5 and 6 to write

$$\Delta_s\bar{z} = \beta_{zw}\mathrm{Var}(w)/\bar{w} = \bar{w}\beta_{zw}\mathrm{Var}(w/\bar{w}). \tag{8}$$

From eqn 3, we have the definition for the average excess in fitness $a_i = w_i/\bar{w} - 1$. Thus, we can expand the expression for the variance in fitness as

$$\mathrm{Var}(w/\bar{w}) = \sum q_i\left(\frac{w_i}{\bar{w}} - 1\right)^2 = \sum q_i a_i^2.$$

From eqn 3, we also have the change in frequency in terms of the average excess, $\Delta q_i = q_i a_i$, and equivalently, $a_i = \Delta q_i/q_i$, thus

$$\begin{aligned}\mathrm{Var}(w/\bar{w}) &= \sum q_i\left(\frac{\Delta q_i}{q_i}\right)^2 \\ &= \sum\left(\frac{\Delta q_i}{\sqrt{q_i}}\right)^2 \\ &= \Delta\hat{\mathbf{q}}\cdot\Delta\hat{\mathbf{q}},\end{aligned}$$

where $\Delta\hat{q}_i = \Delta q_i/\sqrt{q_i}$ is a standardized fluctuation in frequency and $\Delta\hat{\mathbf{q}}$ is the vector of standardized fluctuations. These alternative forms simply express the variance in fitness in different ways. The interesting result follows from the fact that

$$\mathrm{Var}(w/\bar{w}) = \Delta\hat{\mathbf{q}}\cdot\Delta\hat{\mathbf{q}} = F(\Delta\hat{\mathbf{q}})$$

is the Fisher information, F, in the frequency fluctuations, $\Delta\hat{\mathbf{q}}$. Fisher information is a fundamental quantity in information theory, Bayesian analysis, likelihood theory and the informational foundations of statistical inference. Fisher information is a variant form of the more familiar Shannon and Kullback–Leibler information measures, in which the Fisherian form expresses changes in information.

Once again, we have a simple identity. Although it is true that Fisher information is just an algebraic rearrangement of the variance in fitness, some insight may be gained by relating selection to information. The variance form calls to mind a statistical description of selection or a partial description of a probability distribution. The Fisher information form suggests a relation between natural selection and the way in which populations accumulate information (Frank, 2009).

We may now write our fundamental expression for selection as

$$\Delta_s\bar{z} = \bar{w}\beta_{zw}F(\Delta\hat{\mathbf{q}}).$$

We may read this expression for selection as follows: the change in mean character value caused by natural selection, $\Delta_s\bar{z}$, is equal to the total Fisher information in the frequency fluctuations, F, multiplied the scaling β that describes the amount of the potential information that the population captures when expressed in units of phenotypic change. In other words, the distance $\Delta_s\bar{z}$ measures the informational gain by the population caused by natural selection.

The invariances set by this expression may be viewed in different ways. For example, the distance of evolutionary change by selection, $\Delta_s\bar{z}$, is invariant with respect to many different combinations of frequency fluctuations, $\Delta\hat{\mathbf{q}}$, and scalings between phenotype and fitness. Similarly, any transformations of frequency fluctuations that leave the measure of information, $F(\Delta\hat{\mathbf{q}})$, invariant

do not alter the scaled change in phenotype caused by natural selection. The full implications remain to be explored.

Summary of selection identities

The various identities for the part of total evolutionary change caused by selection include

$$\begin{aligned}\Delta_s\bar{z} &= \mathrm{Cov}(w,z)/\bar{w}\\ &= \bar{w}\beta_{zw}\mathrm{Var}(w/\bar{w})\\ &= \boldsymbol{\Delta q}\cdot\mathbf{z}\\ &= \|\boldsymbol{\Delta q}\|\|\mathbf{z}\|\cos\phi\\ &= \bar{w}\beta_{zw}(\boldsymbol{\Delta}\hat{\mathbf{q}}\cdot\boldsymbol{\Delta}\hat{\mathbf{q}})\\ &= \bar{w}\beta_{zw}F(\boldsymbol{\Delta}\hat{\mathbf{q}}).\end{aligned} \tag{9}$$

These forms show the equivalence of the statistical, geometrical and informational expressions for natural selection. These general abstract forms make no assumptions about the nature of phenotypes and the patterns of frequency fluctuations caused by differential fitness. The phenotypes may be squared deviations so that the average is actually a variance, or the product of measurements on different characters leading to measures of association, or any other nonlinear combination of measurements. Thus, there is nothing inherently linear or restrictive about these expressions.

Selection versus evolution

The previous sections discussed the part of evolutionary change caused by selection. The full Price equation (eqn 2) gives a complete and exact expression of total change, repeated here as

$$\Delta\bar{z} = \mathrm{Cov}(w,z)/\bar{w} + \mathrm{E}(w\Delta z)/\bar{w} \tag{10}$$

or in terms of the dot product notation as

$$\Delta\bar{z} = \boldsymbol{\Delta q}\cdot\mathbf{z} + \mathbf{q}'\cdot\boldsymbol{\Delta z}. \tag{11}$$

The full change in the phenotype is the sum of the two terms, which we may express in symbols as

$$\Delta\bar{z} = \Delta_s\bar{z} + \Delta_E\bar{z}.$$

Fisher (1930) called the term $\Delta_E\bar{z}$ the change caused by the environment (Frank & Slatkin, 1992). However, the word *environment* often leads to confusion. The proper interpretation is that $\Delta_E\bar{z}$ encompasses everything not included in the expression for selection. The term is *environmental* only in the sense that it includes all those forces external to the particular definition of the selective forces for a particular problem.

The Δ_E term is sometimes associated with changes in transmission (Frank, 1995, 1997, 2012a; Okasha, 2006). This interpretation arises because $\mathrm{E}(w\Delta z)$ is the fitness-weighted changes in character value between ancestor and descendant. One may think of changes in character values as changes during transmission.

It is important to realize that *everything* truly means every possible force that might arise and that is not accounted for by the particular expression for selection. Lightning may strike. New food sources may appear. The Price equation in its general and abstract form is a mathematical identity–what I previously called a *mathematical tautology* (Frank, 1995).

In applications, one considers how to express $\Delta_E\bar{z}$, or one searches for ways to formulate the problem so that $\Delta_E\bar{z}$ is zero or approximately zero. This article is not about particular applications. Here, I simply note that when one works with Fisher's breeding value as z, near equilibria (fixed points), one typically obtains $\Delta z \to 0$ and thus $\mathrm{E}(w\Delta z) \to 0$. In other cases, the search for a good way to express a problem means finding a form of character measurement that defines z such that characters tend to remain stable over time, so that $\Delta z \to 0$ and thus $\mathrm{E}(w\Delta z) \to 0$. For applications that emphasize calculation of complex dynamics rather than a more abstract conceptual analysis of a problem, methods other than the Price equation often work better.

Abstract properties: recursion and group selection

To iterate is human, to recurse, divine (Coplien, 1998).

Essentially, all modern discussions of multilevel selection and group selection derive from Price (1972a), as developed by Hamilton (1975). Price and Hamilton noted that the Price equation can be expanded recursively to represent nested levels of analysis, for example individuals living in groups.

Start with the basic Price equation as given in eqn 10. The left side is the total change in average phenotype, $\bar{z}$. The second term on the right side includes the terms Δz_i in $\mathrm{E}(w\Delta z) = \sum q_i w_i \Delta z_i$.

Recall that in defining z_i, we specified the meaning of the index i to be any sort of labelling of set members, subject to minimal consistency requirements. We may, for example, label all members of a group by i and measure z_i as some property of the group. If the index i itself represents a set, then we may consider the members of that set. For example, z_{ij} may be the jth member of the ith set, or we may say, the ith group. In the abstract mathematical expression, there is no need to think of the ith group as having any spatial or biological meaning. However, we may consider i as a label for spatially defined groups if we wish to do so.

With i defining a group, we may analyse the selection and evolution of that ith group. The term Δz_i becomes the average change in the z measure for the ith group, composed of members with values z_{ij}. The terms z'_{ij} are the average property values of the descendants of the jth entity in the ith group. The descendant entities that derive from the ith group do not have to form any sort of group or other meaningful structuring, just as the original i labelling does not have to refer to group

structuring in the ancestors. However, we may if we wish consider descendants of i as retaining some sense of the ancestral grouping.

Because z_i represents an averaging over the entities j in the ith group, we are assuming the notational equivalence $\Delta z_i = \Delta\bar{z}_i$. From that point of view, for each group i we may from eqn 10 express the change in the group mean by thinking of each group as a separate set or population, yielding for each i the expression

$$\Delta z_i = \Delta\bar{z}_i = \mathrm{Cov}(w_i, z_i)/\bar{w}_i + \mathrm{E}(w_i\Delta z_i)/\bar{w}_i.$$

We may substitute this expression for each i into the $\mathrm{E}(w\Delta z) = \sum q_i w_i \Delta z_i$ term on the right side of eqn 10. That substitution recursively expands each change in property value, Δz_i, to itself be composed of a selection term and property value change term. For each group, i, we now have expressions for selection within the group, $\mathrm{Cov}(w_i, z_i)/\bar{w}_i$, and average property value change within the group, $\mathrm{E}(w_i\Delta z_i)/\bar{w}_i$. If we write out the full expression for this last term, we obtain

$$\mathrm{E}(w_i\Delta z_i)/\bar{w}_i = \sum_j w_{ij}\Delta z_{ij}/\bar{w}_i.$$

In the term Δz_{ij}, each labelling, j, may itself be a subgroup within the larger grouping represented by i. The recursive nature of the Price equation allows another expansion to the characters z_{ijk} for the kth entity in the jth grouping that is nested in the ith group and so on. Once again, the indexing for levels i, j and k do not have to correspond to any particular structuring, but we may choose to use a structuring if we wish.

One could analyse biological problems of group selection without using the Price equation. Because the Price equation is a mathematical identity, there are always other ways of expressing the same thing. However, in the 1970s, when group selection was a very confused subject, the Price equation's recursive nature and Hamilton's development provided the foundation for subsequent understanding of the topic. All modern conceptual insights about group selection derive from Price's recursive expansion of his abstract expression of selection.

History and alternative expressions of selection

I have emphasized the general and abstract form of the Price equation. That abstract form was first presented rather cryptically by Price (1972a). In that article, Price described the recursive expansion to analyse group selection. Apart from the recursive aspect, the more general abstract properties were hardly mentioned in the study of Price (1972a) and not developed by others until 1995.

While I was writing my history of Price's contributions to evolutionary genetics (Frank, 1995), I found Price's unpublished manuscript *The nature of selection* among W. D. Hamilton's papers. Price's unpublished manuscript gave a very general and abstract scheme for analysing selection in terms of set relations. However, Price did not explicitly connect the abstract set relation scheme to the Price equation or to his earlier publications (Price, 1970, 1972a).

I had *The nature of selection* published posthumously as Price (1995). In my own article, I explicitly developed the general interpretation of the Price equation as the formal abstract expression of the relation between two sets (Frank, 1995).

Price (1970) wrote an earlier article in which he presented a covariance selection equation that emphasized the connection to classical models of population genetics and gene frequency change. That earlier covariance form lacks the abstract set interpretation and generally has narrower scope. Preceding Price, Robertson (1966) and Li (1967) also presented selection equations that are similar to Price's (1970) covariance expression. Robertson's covariance form itself arises from classical quantitative genetics and the breeder's equation, ultimately deriving from the foundations of quantitative genetics established by Fisher (1918). Li's form presents a covariance type of expression for classical population genetic models of gene frequency change.

One cannot understand the current literature without a clear sense of this history. Almost all applications of the Price equation to kin and group selection, and to other problems of evolutionary analysis, derive from either the classical expressions of quantitative genetics (Robertson, 1966) or classical expressions of population genetics (Li, 1967).

In the light of this history, criticisms can be confusing with regard to the ways in which the Price equation is commonly used. For example, in applications to kin or group selection, the Price equation mainly serves to package the notation for the Robertson form of quantitative genetic analysis or the Li form of population genetic analysis. The Price equation packaging brings no extra assumptions. In some applications, critics may believe that the particular analysis lacks enough assumptions to attain a desired level of specificity. One can, of course, easily add more assumptions, at the expense of reduced generality.

The following sections briefly describe some alternative forms of the Price equation and the associated history. That history helps to place criticisms of the Price equation and its applications into clearer light.

Quantitative genetics and the breeder's equation

Fisher (1918) established the modern theory of quantitative genetics, following the early work of Galton, Pearson, Weldon, Yule and others. The equations of selection in quantitative genetics and animal breeding arose from that foundation. Many modern applications of the Price equation to particular problems follow this tradition of quantitative genetics. A criticism of these Price equation applications is a criticism of the central approach of evolutionary quantitative genetics. Such

criticisms may be valid for certain applications, but they must be evaluated in the broader context of quantitative genetics theory. This section shows the relation between quantitative genetics and a commonly applied form of the Price equation (Rice, 2004).

Evolutionary aspects of quantitative genetics developed from the breeder's equation

$$R = Sh^2,$$

in which the response to selection, R, equals the selection differential, S, multiplied by the heritability, h^2. The separation of selection and transmission is the key to the breeder's equation and to quantitative genetics theory.

The covariance term of the Price equation is equivalent to the selection differential, S, when one interprets the meaning of *fitness* and *descendants* in a particular way. Suppose that we label each potential parent in the ancestral population of size N with the index, i. The initial weighting of each parent in the ancestral population is $q_i = 1/N$. Assign to each potential parent a weighting with respect to breeding contribution, $q_i' = q_i w_i$, with fitnesses standardized so that $\bar{w} = 1$ and the w_i are relative fitnesses.

With this set-up, ancestors are the initial population of potential parents, each weighted equally, and descendants are the same population of parents, weighted by their breeding contribution. The character value for each individual remains unchanged between the ancestor and descendant labellings. These assumptions lead to $\Delta\bar{z}^* = \mathrm{Cov}(w,z)$, the change in the average character value between the breeding population and the initial population. That difference is defined as S, the selection differential.

To analyse the fraction of the selection differential transmitted to offspring, classical quantitative genetics follows Fisher (1918) to separate the character value as $z = g + \epsilon$, with a transmissible genetic component, g, and a component that is not transmitted, which we may call the environmental or unexplained component, ϵ. Following standard regression theory for this sort of expression, $\bar{\epsilon} = 0$.

For a parent with $z = g + \epsilon$, the average character value contribution ascribed to the parent among its descendants is $z' = g$, following the idea that g represents the component of the parental character that is transmitted to offspring. If we assume that the only fluctuations of average character value in offspring are caused by the transmissible component that comes from parents, then the genetic component measured by g is sufficient to explain expected offspring character values. Thus, $\Delta z = z' - z = -\epsilon$, and $\mathrm{E}(w\Delta z) = -\mathrm{Cov}(w,\epsilon)$.

Substituting into the full Price equation from eqn 2 and assuming $\bar{w} = 1$ so that all fitnesses are normalized

$$\begin{aligned}\Delta\bar{z} &= \mathrm{Cov}(w,z) + \mathrm{E}(w\Delta z)\\ &= \mathrm{Cov}(w,g) + \mathrm{Cov}(w,\epsilon) - \mathrm{Cov}(w,\epsilon)\\ &= \mathrm{Cov}(w,g). \qquad (12)\end{aligned}$$

The expression $\Delta\bar{z} = \mathrm{Cov}(w,g)$ was first emphasized by Robertson (1966) and is sometimes called Robertson's secondary theorem of natural selection. Robertson's expression summarizes the foundational principles of quantitative genetics, as conceived by Fisher (1918) and developed over the past century (Falconer & Mackay, 1996; Lynch & Walsh, 1998; Hartl, 2006).

It is commonly noted that Robertson's theorem is related to the classic breeder's equation. In particular,

$$R = \Delta\bar{z} = \mathrm{Cov}(w,g) = \mathrm{Cov}(w,z)h^2 = Sh^2,$$

where R is the response to selection, $S = \mathrm{Cov}(w,z)$ is the selection differential and $h^2 = \mathrm{Var}(g)/\mathrm{Var}(z)$ is a form of heritability, a measure of the transmissible genetic component. Additional details and assumptions can be found in several articles and texts (Crow & Nagylaki, 1976; Frank, 1997; Rice, 2004).

Population genetics and the covariance expression

Price (1970) expressed his original formulation in terms of gene frequency change and classical population genetics, rather than the abstract set relations that I have emphasized. At that time, it seems likely that Price already had the broader, more abstract theory in hand and was presenting the population genetics form because of its potential applications. The article begins

> This is a preliminary communication describing applications to genetical selection of a new mathematical treatment of selection in general.
> Gene frequency change is the basic event in biological evolution. The following equation… which gives frequency change under selection from one generation to the next for a single gene or for any linear function of any number of genes at any number of loci, holds for any sort of dominance or epistasis, for sexual or asexual reproduction, for random or nonrandom mating, for diploid, haploid or polyploid species, and even for imaginary species with more than two sexes…

Using my notation, Price writes the basic covariance form

$$\Delta P = \mathrm{Cov}(w,p)/\bar{w} = \beta_{wp}\mathrm{Var}(p)/\bar{w}. \qquad (13)$$

In a simple application, p could be interpreted as gene frequency at a single diploid locus with two alleles. Then, $P = \bar{p}$ is the gene frequency in the population, and β_{wp} is the regression of individual fitness on individual gene frequency, in which the individual gene frequency is either 0, 1/2 or 1 for an individual with 0, 1 or 2 copies of the allele of interest. Li (1967) gave an identical gene frequency expression in his eqn 4.

In more general applications, one can study a p-score that summarizes the number of copies of various alleles present in an individual or in whatever entities are being tracked. In classical population genetics, the p-score would be, in Price's words above, 'any linear function of any number of genes at any number of loci'. Here, *linearity* means that p is essentially a counting of presence versus absence of various things within the ith entity.

Such counting does not preclude nonlinear interactions between alleles or those things being counted with respect to phenotype, which is why Price said that the expression holds for any form of dominance or epistasis.

Hamilton (1970) used Price's gene frequency form in his first clear derivations of the direct and the inclusive fitness models of kin selection theory. Most early applications of the Price equation used this gene frequency interpretation.

Price (1970) emphasized that the value of eqn 13 arises from its benefits for qualitative reasoning rather than calculation. The necessary assumptions can be seen from the form given by Price, which is always exact, here written in my notation

$$\Delta P = \mathrm{Cov}(w, p)/\bar{w} + \mathrm{E}(w\Delta p)/\bar{w},$$

where Δp is interpreted as the change in state between parental gene frequency for the ith entity and the average gene frequency for the part of descendants derived from the ith entity.

In practice, $\Delta p = 0$ usually means Mendelian segregation, no biased mutation and no sampling biases associated with drift. Most population genetics theory of traits such as social behaviour typically make those assumptions, so that eqn 13 is sufficient with respect to analysing change in gene frequency or in p-scores (Grafen, 1984). However, the direction of change in gene frequency or p-score is not sufficient to predict the direction of change in phenotype. To associate the direction of change in p-score with the direction of change in phenotype, one must make the assumption that phenotype changes monotonically with p-score. Such monotonicity is a strong assumption, which is not always met. For that reason, p-score models sometimes buy simplicity at a rather high cost. In other applications, monotonicity is a reasonable assumption, and the p-score models provide a very simple and powerful approach to understanding the direction of evolutionary change.

The costs and benefits of the p-score model are not particular to the Price equation. Any analysis based on the same assumptions has the same limitations. The Price equation provides a concise and elegant way to explore the consequences when certain simplifying assumptions can reasonably be applied to a particular problem.

Alternative forms or interpretations of the full equation

The full Price equation partitions total evolutionary change into components. Many alternative partitions exist. A partition provides value if it improves conceptual clarity or eases calculation.

Which partitions are better than others? *Better* is always partly subjective. What may seem hard for me may appear easy to you. Nonetheless, it would be a mistake to suggest that all differences are purely subjective. Some forms are surely better than others for particular problems, even if *better* remains hard to quantify. As Russell (1958, p. 14) said in another context, 'All such conventions are equally legitimate, though not all are equally convenient'.

Many partitions of evolutionary change include some aspect of selection and some aspect of property or transmission change. Most of those variants arise by minor rearrangements or extensions of the basic Price expression. A few examples follow.

Contextual analysis

Heisler & Damuth (1987) introduced the phrase *contextual analysis* to the evolutionary literature. Contextual analysis is a form of path analysis, which partitions causes by statistical regression models. Path analysis has been used throughout the history of genetics (Li, 1975). It is a useful approach whenever one wishes to partition variation with respect to candidate causes. The widely used method of Lande & Arnold (1983) to analyse selection is a particular form of path analysis.

Okasha (2006) argued that contextual analysis is an alternative to the Price equation. To develop a simple example, let us work with just the selection part of the Price equation

$$\bar{w}\Delta\bar{z} = \mathrm{Cov}(w, z).$$

A path (contextual) analysis refines this expression by partitioning the causes of fitness with a regression equation. Suppose we express fitness as depending on two predictors: the focal character that we are studying, z, and another character, y. Then we can write fitness as

$$w = \beta_{wz}z + \beta_{wy}y + \epsilon$$

in which the β terms are partial regressions of fitness on each character, and ϵ is the unexplained residual of fitness. Substituting into the Price equation, we get the sort of expression made popular by Lande & Arnold (1983)

$$\bar{w}\Delta\bar{z} = \beta_{wz}\mathrm{Var}(z) + \beta_{wy}\mathrm{Cov}(y, z).$$

If the partitioning of fitness into causes is done in a useful way, this type of path analysis can provide significant insight. I based my own studies of natural selection and social evolution on this approach (Frank, 1997, 1998).

Authors such as Okasha (2006) consider the partitioning of fitness into distinct causes as an alternative to the Price equation. If one thinks of the character z in $\mathrm{Cov}(w,z)$ as a complete causal explanation for fitness, then a partition into separate causes y and z does indeed lead to a different causal understanding of fitness. In that regard, the Price equation and path analysis lead to different causal perspectives.

One can find articles that use the Price equation and interpret z as a lone cause of fitness (see Okasha, 2006). Thus, if one equates those specific applications with the

general notion of *the Price equation*, then one can say that path or contextual analysis provides a significantly different perspective from the Price equation. To me, that seems like a socially constructed notion of logic and mathematics. If someone has applied an abstract truth in a specific way, and one can find an alternative method for the same specific application that seems more appealing, then one can say that the alternative method is superior to the general abstract truth.

The abstract Price equation does not compel one to interpret z strictly as a single-cause explanation. Rather, in the general expression, z should always be interpreted as an abstract placeholder. Path (contextual) analysis follows as a natural extension of the Price equation, in which one makes specific models of fitness expressed by regression. It does not make sense to discuss the Price equation and path analysis as alternatives.

Alternative partitions of selection and transmission

In the standard form of the Price equation, the fitness term, w, appears in both components

$$\bar{w}\Delta\bar{z} = \mathrm{Cov}(w, z) + \mathrm{E}(w\Delta z).$$

Frank (1997, 1998) derived an alternative expression

$$\begin{aligned}\Delta\bar{z} &= \sum q_i' z_i' - \sum q_i z_i \\ &= \sum q_i (w_i/\bar{w}) z_i' - \sum q_i z_i \\ &= \sum q_i (w_i/\bar{w}) z_i' - \sum q_i z_i' + \sum q_i z_i' - \sum q_i z_i \\ &= \sum q_i (w_i/\bar{w} - 1) z_i' + \sum q_i (z_i' - z_i) \\ &= \mathrm{Cov}(w, z')/\bar{w} + \mathrm{E}(\Delta z). \end{aligned} \quad (14)$$

This form sometimes provides an easier method to calculate effects. For example, the second term now expresses the average change in phenotype between parent and offspring without weighting by fitness effects. A biased mutational process would be easy to calculate with this expression – one only needs to know about the mutation process to calculate the outcome. The new covariance term can be partitioned into meaningful components with minor assumptions (Frank, 1997, p. 1721), yielding

$$\mathrm{Cov}(w, z') = \mathrm{Cov}(w, z)\beta_{z'z},$$

where $\beta_{z'z}$ is usually interpreted as the offspring–parent regression, which is a type of heritability. Thus, we may combine selection with the heritability component of transmission into the covariance term, with the second term containing only a fitness-independent measure of change during transmission.

Okasha (2006) strongly favoured the alternative partition for the Price equation in eqn 14, because it separates all fitness effects in the first term from a pure transmission interpretation of the second term. In my view, there are costs and benefits for the standard Price equation expression compared with eqn 14. One gains by having both and using the particular form that fits a particular problem.

For example, the term $\mathrm{E}(\Delta z)$ is useful when one has to calculate the effects of a biased mutational process that operates independently of fitness. Alternatively, suppose most individuals have unbiased transmission, such that $\Delta z = 0$, whereas very sick individuals do not reproduce but, if they were to reproduce, would have a very biased transmission process. Then $\mathrm{E}(\Delta z)$ differs significantly from zero, because the sick, nonreproducing individuals appear in this term equally with the reproducing population. However, the actual transmission bias that occurs in the population would be zero, $\mathrm{E}(w\Delta z) = 0$, because all reproducing individuals have nonbiased transmission.

Both the standard Price form and the alternative in eqn 14 can be useful. Different scenarios favour different ways of expressing problems. I cannot understand why one would adopt an a priori position that unduly limits one's perspective.

Extended set mapping expression

The Price equation's power arises from its abstraction of selection in terms of mapping relations between sets (Frank, 1995; Price, 1995). Although the Price equation is widely cited in the literature, almost no work has developed the set mapping formalism beyond the description given in the initial publications. I know of only one article.

Kerr & Godfrey-Smith (2009) noted that, in the original Price formulation, every descendant must derive from one or more ancestors. There is no natural way for novel entities to appear. In applications, new entities could arise by immigration from outside the system or, in a cultural interpretation, by de novo generation of an idea or behaviour.

Kerr & Godfrey-Smith (2009) present an extended expression to handle unconnected descendants. Their formulation depends on making explicit the connection number between each individual ancestor and each individual descendant, rather than using the fitnesses of types. Some descendants may have zero connections.

With an explicit description of connections, an extended Price equation follows. The two core components of covariance for selection and expected change for transmission occur, plus a new factor to account for novel descendants unconnected to ancestors.

The notation in Kerr & Godfrey-Smith (2009) is complex, so I do not repeat it here. Instead, I show a simplified version. Suppose that a fraction p of the descendants are unconnected to ancestors. Then, we can write the average trait value among descendants as

$$\bar{z}' = p\sum \alpha_j z_j^* + (1-p)\sum q_i' z_i',$$

where z_j^* is the phenotype for the jth member of the descendant population that is unconnected to ancestors and α_j is the frequency of each unconnected type, with

$\sum \alpha_j = 1$. Given those definitions, we can proceed with the usual Price equation expression

$$\begin{aligned}\Delta\bar{z} &= \bar{z}' - \bar{z} \\ &= p\sum \alpha_j z_j^* + (1-p)\sum q_i' z_i' - (p+1-p)\sum q_i z_i \\ &= (1-p)\Big(\sum q_i' z_i' - \sum q_i z_i\Big) + p\Big(\sum \alpha_j z_j^* - \sum q_i z_i\Big).\end{aligned}$$

Note that the term weighted by $1-p$ leads to the standard form of the Price equation, so we can write

$$\begin{aligned}\Delta\bar{z} &= (1-p)(\mathrm{Cov}(w,z) + \mathrm{E}(w\Delta z))/\bar{w} \\ &\quad + p\Big(\sum \alpha_j z_j^* - \sum q_i z_i\Big) \\ &= (1-p)(\mathrm{Cov}(w,z) + \mathrm{E}(w\Delta z))/\bar{w} + p(\bar{z}^* - \bar{z}).\end{aligned}$$

In the component weighted by p, no connections exist between the descendant z_j^* and a member of the ancestral population. Thus, we have no basis to relate those terms to fitness, transmission or property change. Kerr & Godfrey-Smith (2009) use an alternative notation that associates all entities with their number of connections, including those with zero. The outcome is an extended set mapping theory for evolutionary change. The main concepts and the value of the approach are best explained by the application presented in the next section.

Gains and losses in descendants and ancestors

Fox & Kerr (2012) analyse changes in ecosystem function by modifying the method of Kerr & Godfrey-Smith (2009). They measure ecosystem function by summing the functional contribution of each species present in an ecosystem. To compare ecosystems, they consider an initial site and a second site. When comparing ecosystems, the notion of ancestors and descendants may not make sense. Instead, one appeals to the more general set mapping relations of the Price equation.

Assume that there is an initial site with total function $T = \sum z_i$, where z_i is the function of the ith species. At the initial site, there are s different species; thus, we may also express the total as $T = s\bar{z}$, where $\bar{z}$ is the average function per species. At a second site, total function is $T' = \sum z_j'$, with s' different species in the summation and $T' = s'\bar{z}'$. Let the number of species in common between the sites be s_c. Thus, the initial site has $S = s - s_c$ unique species, and the second site has $S' = s' - s_c$ unique species.

Fox & Kerr (2012) write the change in total ecosystem function as

$$\begin{aligned}\Delta T = T' - T = {} & s'\bar{z}' - s\bar{z} \\ = {} & (s' - s_c)\bar{z}' - (s - s_c)\bar{z} + s_c(\bar{z}' - \bar{z}) \\ = {} & S'\bar{z}' - S\bar{z} + s_c(\Delta\bar{z}).\end{aligned}$$

The term $S'\bar{z}'$ represents the change in function that caused the gain of an average species, in which S' is the number of newly added species, and $\bar{z}'$ is the average function per species. Fox & Kerr (2012) suggest that a randomly added species would be expected to function as an average species, and so interpret this term as the contribution of random species gain. The term $S\bar{z}$ is interpreted similarly as random species loss with respect to the S unique species in the first ecosystem not present in the second ecosystem.

Fox & Kerr (2012) partition the term $s_c(\Delta\bar{z})$ into three components of species function: deviation from the average for species gained at the second site; deviation from the average for species lost from the first site; and the changes in function for those species in common between sites.

The point here concerns the approach rather than the theory of ecosystem function. To analyse changes between two sets, one often benefits by an explicit decomposition of the relations between the two sets. The original Price equation is one sort of decomposition, based on tracing the ways in which descendants derive from and change with respect to ancestors. Fox & Kerr (2012) extend the decomposition of change by set mapping to include specific components that make sense in the context of changes in ecosystem function.

More work on the mathematics of set mapping and decomposition would be very valuable. The Price equation and the extensions by Kerr, Godfrey-Smith and Fox show the potential for thinking carefully about the abstract components of change between sets and how to apply that abstract understanding to particular problems.

Other examples

No clear guidelines determine what constitutes an extension to the Price equation. From a broad perspective, many different partitions of total change have similarities, because they separate something like selection from other forces that alter the similarity between populations.

For example, the stochastic effects of sampling and drift create a distribution of descendant phenotypes around the ancestral mean. In the classical Price formulation, there is only the single realization of the actual descendants. A stochastic version analyses a collection of possible descendant sets over some probability distribution and a mapping from the ancestor set to each possible realization of the descendant set.

In other cases, partitions will split components more finely or add new components not in Price's formulation. I do not have space to review every partition of total change and consider how each may be related to Price's formulation. I list a few examples here.

Grafen (1999) and Rice (2008) developed stochastic approaches. Grafen (2007) based a long-term project on interpretations and extensions of the Price equation. Page & Nowak (2002) related the Price equation to various other evolutionary analyses, providing some minor extensions. Wolf *et al.* (1998), Bijma & Wade (2008), and many others developed extended partitions by splitting causes with regression or similar methods such as path analysis.

Various forms of the Price equation have been applied in economic theory (Andersen, 2004).

Difficulties with various critiques of the Price equation

> A reliable way to make people believe in falsehoods is frequent repetition, because familiarity is not easily distinguished from truth (Kahneman, 2011, p. 62).

One must distinguish the full, exact Price equation from various derived forms used in applications. The derived forms always make additional assumptions or express approximate relations (Frank, 1997). Each assumption increases specificity and reduces generality in relation to particular goals.

Critiques of the Price equation rarely distinguish the costs and benefits of particular assumptions in relation to particular goals. I use van Veelen's recent series of papers as a proxy for those critiques. That series repeats some of the common misunderstandings and adds some new ones. Nowak recently repeated van Veelen's critique as the basis for his commentary on the Price equation (van Veelen, 2005; Nowak *et al.*, 2010; van Veelen *et al.*, 2010; Nowak & Highfield, 2011; van Veelen, 2011; van Veelen *et al.*, 2012).

Dynamic sufficiency

The Price equation describes the change in some measurement, expressed as $\Delta\bar{z}$. Change is calculated with respect to particular mapping relations between ancestor and descendant populations. We can think of the mappings and the beginning value of $\bar{z}$ as the initial conditions or inputs and $\Delta\bar{z}$ as the output.

The output, $\bar{z}' = \bar{z} + \Delta\bar{z}$, does not provide enough information to iterate the calculation of change to get another value of $\Delta\bar{z}$ starting with $\bar{z}'$. We would also need the mapping relations between the new descendant population and its subsequent descendants. That information is not part of the initial input. Thus, we cannot study the dynamics of change over time without additional information.

This limitation with regard to repeated iteration is called a lack of *dynamic sufficiency* (Lewontin, 1974). Confusion about the nature of dynamic sufficiency in relation to the Price equation has been common in the literature. In Frank (1995, pp. 378–379), I wrote

> It is not true, however, that dynamic sufficiency is a property that can be ascribed to the Price Equation—this equation is simply a mathematical tautology for the relationship among certain quantities of populations. Instead, dynamic sufficiency is a property of the assumptions and information provided in a particular problem, or added by additional assumptions contained within numerical techniques such as diffusion analysis or applied quantitative genetics. … What problems can the Price equation solve that cannot be solved by other methods? The answer is, of course, none, because the Price Equation is derived from, and is no more than, a set of notational conventions. It is a mathematical tautology.

I showed how the Price equation helps to define the necessary conditions for dynamic sufficiency. Once again, the Price equation proves valuable for clarifying the abstract structure of evolutionary analysis.

Compare my statement with that of van Veelen *et al.* (2012)

> Dynamic insufficiency is regularly mentioned as a drawback of the Price equation (see for example Frank, 1995; Rice, 2004). We think that this is not an entirely accurate description of the problem. We would like to argue that the perception of dynamic insufficiency is a symptom of the fundamental problem with the Price equation, and not just a drawback of an otherwise fine way to describe evolution. To begin with, it is important to realize that the Price equation itself, by its very nature, cannot be dynamically sufficient or insufficient. The Price equation is just an identity. If we are given a list of numbers that represent a transition from one generation to the next, then we can fill in those numbers in both the right and the left hand side of the Price equation. The fact that it is an identity guarantees that the numbers that appear on both sides of the equality sign are the same. There is nothing dynamically sufficient or insufficient about that (this point is also made by Gardner et al., 2007, p. 209). A model, on the other hand, can be dynamically sufficient or insufficient.

This quote from van Veelen *et al.* (2012) demonstrates an interesting approach to scholarship. They first cite Frank as stating that dynamic insufficiency is a drawback of the Price equation. They then disagree with that point of view and present as their own interpretation an argument that is nearly identical in concept and phrasing to my own statement in the very paper that they cited as the foundation for their disagreement.

In this case, I think it is important to clarify the concepts and history, because influential and widely cited authors, such as Nowak, are using van Veelen's articles as the basis for their own critiques of the Price equation and approaches to fundamental issues of evolutionary analysis.

With regard to dynamics, any analysis achieves the same dynamic status given the same underlying assumptions. The Price equation, when used with the same underlying assumptions as population genetics, has the same attributes of dynamic sufficiency as population genetics.

Interpretation of covariance

van Veelen *et al.* (2012) claim that

> Maybe the most unfortunate thing about the Price equation is that the term on the right hand side is denoted as a covariance, even though it is not. The equation thereby turns into something that can easily set us off in the wrong direction, because it now resembles equations as they feature in other sciences, where probabilistic models are used that do use actual covariances.

One can see the covariance expression in the standard form of the Price equation given in eqn 2. In the Price equation, the covariance is measured with respect to the total population; in other words, it expresses the association over all members of the population. In many statistical applications, one only has data on a subset of the full population, that subset forming a sample. It is important to distinguish between population measures and sample measures, because they refer to different things.

Price (1972a, p. 485) made clear that his equation is about total change in entire populations, so the covariance is interpreted as a population measure

> [W]e will be concerned with population functions and make no use of sample functions, hence we will not observe notational conventions for distinguishing population and sample variables and functions.

In addition to population and sample measures, covariance also arises in mathematical models of process. Suppose, for example, that I develop a model in which random processes influence fitness and random processes influence phenotype. If the random fluctuations in fitness and the random fluctuations in phenotype are associated, the random variables of fitness and phenotype would covary. All of these different interpretations of covariance are legitimate; they simply reflect different situations.

Discussion

In Frank (1995), I wrote: 'What problems can the Price equation solve that cannot be solved by other methods? The answer is, of course, none, because the Price Equation is derived from, and is no more than, a set of notational conventions. It is a mathematical tautology'.

Nowak & Highfield (2011) and van Veelen *et al.* (2012) emphasize the same point in their critique of the Price equation, although they present the argument as a novel insight without attribution. Given that the Price equation is a set of notational conventions, it cannot uniquely specify any predictions or insights. A particular set of assumptions leads to the same predictions, no matter what notational conventions one uses. The Price equation is a tool that sometimes helps in analysis or in seeing general connections between apparently disparate ideas. For many problems, the Price equation provides no value, because it is the wrong tool for the job.

If the Price equation is just an equivalence, or tautology, then why am I enthusiastic about it? Mathematics is, in its essence, about equivalences, as expressed beautifully in the epigraph from Mazur. Not all equivalences are interesting or useful, but some are, just as not all mathematical expressions are interesting or useful, but some are.

That leads us to the question of how we might know whether the Price equation is truly useful or a mere identity? It is not always easy to say exactly what makes an abstract mathematical equivalence interesting or useful. However, given the controversy over the Price equation, we should try. Because there is no single answer, or even a truly unique and unambiguous question, the problem remains open. I list a few potential factors.

'[A] good notation has a subtlety and suggestiveness which at times make it seem almost like a live teacher' (Russell, 1922, pp. 17–18). Much of creativity and understanding comes from seeing previously hidden associations. The tools and forms of expression that we use play a strong role in suggesting connections and are inseparable from cognition (Kahneman, 2011). Equivalences and alternative notations are important.

The various forms of the covariance component from the Price equation given in eqn 9 show the equivalence of the statistical, geometrical and informational expressions for natural selection. The recursive form of the full Price equation provides the foundation for all modern studies of group selection and multilevel analysis. The Price equation helped in discovering those various connections, although there are many other ways in which to derive the same relations.

Hardy (1967) also emphasized the importance of seeing new connections between apparently disparate ideas:

> We may say, roughly, that a mathematical idea is 'significant' if it can be connected, in a natural and illuminating way, with a large complex of other mathematical ideas. Thus a serious mathematical theorem, a theorem which connects significant ideas, is likely to lead to important advances in mathematics itself and even in other sciences.

What sort of connections? One type concerns the invariances discovered or illuminated by the Price equation. I discussed some of those invariances in an earlier section, particularly the information theory interpretation of natural selection through the measure of Fisher information (Frank, 2009). Fisher's fundamental theorem of natural selection is a similar sort of invariance (Frank, 2012b). Kin selection theory derives much of its power by identifying an invariant informational quantity sufficient to unify a wide variety of seemingly disparate processes (Frank, 1998, Chapter 6). The interpretation of kin selection as an informational invariance has not been fully developed and remains an open problem.

Invariances provide the foundation of scientific understanding: 'It is only slightly overstating the case to say that physics is the study of symmetry' (Anderson, 1972). *Invariance* and *symmetry* mean the same thing (Weyl, 1983). Feynman (1967) emphasized that invariance is *The Character of Physical Law.* The commonly observed patterns of probability can be unified by the study of invariance and its association with measurement (Frank & Smith, 2010, 2011). There has been little effort in biology to pursue similar understanding of invariance and measurement (Frank, 2011; Houle *et al.*, 2011).

Price argued for the great value of abstraction, in the sense of the epigraph from Mazur. In Price (1995)

> [D]espite the pervading importance of selection in science and life, there has been no abstraction and generalization from genetical selection to obtain a general selection theory and general selection mathematics. Instead, particular selection problems are treated in ways appropriate to particular fields of science. Thus one might say that 'selection theory' is a theory waiting to be born—much as communication theory was 50 years ago. Probably the main lack that has been holding back any development of a general selection theory is lack of a clear concept of the general nature or meaning of 'selection'.

This article has been about the Price equation in relation to its abstract properties and its connections to various topics, such as information or fundamental invariances. Some readers may feel that those aspects of abstraction and invariance are nice, but far from daily work in biology. What of the many applications of the Price equation to kin or group selection? Do those applications hold up? How much value has been added?

Because the Price equation is a tool, one can always arrive at the same result by other methods. How well the Price equation works depends partly on the goal and partly on the fit of the tool to the problem. There is inevitably a strongly subjective aspect to deciding about how well a tool works. Nonetheless, hammers truly are good for nails and bad for screws. For valuing tools, there is a certain component that should be open to agreement. For example, the Robertson (1966) form of the Price equation is widely regarded as the foundational method for analysing models of evolutionary quantitative genetics. However, not all problems in quantitative genetics are best studied with the Robertson–Price equation. And not all problems in social evolution benefit from a Price equation approach.

The Price equation or descendant methods have led to many useful models for kin selection (Frank, 1998). The most powerful follow a path analysis decomposition of causes or use a simple maximization method to analyse easily what would otherwise have been difficult. I will return to those applications in subsequent articles.

Acknowledgments

I thank R.M. Bush and W.J. Ewens for helpful comments. My research is supported by National Science Foundation grant EF-0822399, National Institute of General Medical Sciences MIDAS Program grant U01-GM-76499 and a grant from the James S. McDonnell Foundation.

References

Alizon, S. 2009. The Price equation framework to study disease within-host evolution. *J. Evol. Biol.* **22**: 1123–1132.

Andersen, E.S. 2004. Population thinking, Price's equation and the analysis of economic evolution. *Evol. Inst. Econ. Rev.* **1**: 127–148.

Anderson, P. 1972. More is different. *Science* **177**: 393–396.

Bijma, P. & Wade, M.J. 2008. The joint effects of kin, multilevel selection and indirect genetic effects on response to genetic selection. *J. Evol. Biol.* **21**: 1175–1188.

Boulding, K.E. 1956. General systems theory: the skeleton of science. *Manage. Sci.* **2**: 197–208.

Charlesworth, B. & Charlesworth, D. 2010. *Elements of Evolutionary Genetics*. Roberts & Company, Greenwood Village, CO.

Coplien, J.O. 1998. To iterate is human, to recurse, divine. *C++ Rep.* **10**: 43–51.

Crow, J.F. & Kimura, M. 1970. *An Introduction to Population Genetics Theory*. Burgess, Minneapolis, Minnesota.

Crow, J.F. & Nagylaki, T. 1976. The rate of change of a character correlated with fitness. *Am. Nat.* **110**: 207–213.

Day, T. & Gandon, S. 2006. Insights from Price's equation into evolutionary epidemiology. In: *Disease Evolution: Models, Concepts, and Data Analysis* (Z. Feng, U. Dieckmann & S. Levin, eds), pp. 23–44. American Mathematical Society, Washington, DC.

Ewens, W.J. 1989. An interpretation and proof of the fundamental theorem of natural selection. *Theor. Popul. Biol.* **36**: 167–180.

Ewens, W.J. 1992. An optimizing principle of natural selection in evolutionary population genetics. *Theor. Popul. Biol.* **42**: 333–346.

Falconer, D.S. & Mackay, T.F.C. 1996. *Introduction to Quantitative Genetics*, 4th edn. Longman, Essex, UK.

Feynman, R.P. 1967. *The Character of Physical Law*. MIT Press, Cambridge, MA.

Fisher, R.A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edinb.* **52**: 399–433.

Fisher, R.A. 1930. *The Genetical Theory of Natural Selection*. Clarendon, Oxford.

Fisher, R.A. 1941. Average excess and average effect of a gene substitution. *Ann. Eugen.* **11**: 53–63.

Fox, J.W. 2006. Using the Price Equation to partition the effects of biodiversity loss on ecosystem function. *Ecology* **87**: 2687–2696.

Fox, J.W. & Kerr, B. 2012. Analyzing the effects of species gain and loss on ecosystem function using the extended Price equation partition. *Oikos* **121**: 290–298.

Frank, S.A. 1995. George Price's contributions to evolutionary genetics. *J. Theor. Biol.* **175**: 373–388.

Frank, S.A. 1997. The Price equation, Fisher's fundamental theorem, kin selection, and causal analysis. *Evolution* **51**: 1712–1729.

Frank, S.A. 1998. *Foundations of Social Evolution*. Princeton University Press, Princeton, NJ.

Frank, S.A. 2009. Natural selection maximizes Fisher information. *J. Evol. Biol.* **22**: 231–244.

Frank, S.A. 2011. Measurement scale in maximum entropy models of species abundance. *J. Evol. Biol.* **24**: 485–496.

Frank, S.A. 2012a. Natural selection. III. Selection versus transmission and the levels of selection. *J. Evol. Biol.* **25**: 227–243.

Frank, S.A. 2012b. Wright's adaptive landscape versus Fisher's fundamental theorem. In: *The Adaptive Landscape in Evolutionary Biology* (E. Svensson & R. Calsbeek, eds). Oxford University Press, New York (in press). Available from http://arxiv.org/abs/1102.3709v1.

Frank, S.A. & Slatkin, M. 1990. The distribution of allelic effects under mutation and selection. *Genet. Res.* **55**: 111–117.
Frank, S.A. & Slatkin, M. 1992. Fisher's fundamental theorem of natural selection. *Trends Ecol. Evol.* **7**: 92–95.
Frank, S.A. & Smith, E. 2010. Measurement invariance, entropy, and probability. *Entropy* **12**: 289–303.
Frank, S.A. & Smith, E. 2011. A simple derivation and classification of common probability distributions based on information symmetry and measurement scale. *J. Evol. Biol.* **24**: 469–484.
Frieden, B.R. 2004. *Science from Fisher Information: A Unification*. Cambridge University Press, Cambridge, UK.
Frieden, B.R., Plastino, A. & Soffer, B.H. 2001. Population genetics from an information perspective. *J. Theor. Biol.* **208**: 49–64.
Gardner, A., West, S.A. & Barton, N.H. 2007. The relation between multilocus population genetics and social evolution theory. *Am. Nat.* **169**: 207–226.
Gardner, A. 2008. The Price equation. *Curr. Biol.* **18**: R198–R202.
Grafen, A. 1984. Natural selection, kin selection and group selection. In: *Behavioural ecology* (J.R. Krebs & N.B. Davies, eds), pp. 62–84. Blackwell Scientific Publications, Oxford.
Grafen, A. 1999. Formal Darwinism, the individual-as-maximizing-agent analogy and bet-hedging. *Proc. R. Soc. Lond. B* **266**: 799–803.
Grafen, A. 2002. A first formal link between the Price equation and an optimization program. *J. Theor. Biol.* **217**: 75–91.
Grafen, A. 2007. The formal Darwinism project: a mid-term report. *J. Evol. Biol.* **20**: 1243–1254.
Hamilton, W.D. 1970. Selfish and spiteful behaviour in an evolutionary model. *Nature* **228**: 1218–1220.
Hamilton, W.D. 1975. Innate social aptitudes of man: an approach from evolutionary genetics. In: *Biosocial Anthropology* (R. Fox, ed.), pp. 133–155. Wiley, New York.
Hardy, G.H. 1967. *A Mathematician's Apology*. Cambridge University Press, Cambridge, UK.
Harman, O.S. 2010. *The Price of Altruism*. Norton, New York.
Hartl, D.L. 2006. *Principles of Population Genetics*, 4th edn. Sinauer, Sunderland, MA.
Heisler, I.L. & Damuth, J. 1987. A method for analyzing selection in hierarchically structured populations. *Am. Nat.* **130**: 582–602.
Houle, D., Pélabon, C., Wagner, G.P. & Hansen, T.F. 2011. Measurement and meaning in biology. *Q. Rev. Biol.* **86**: 3–34.
Kahneman, D. 2011. *Thinking, Fast and Slow*. Farrar Straus & Giroux, New York.
Kerr, B. & Godfrey-Smith, P. 2009. Generalization of the Price equation for evolutionary change. *Evolution* **63**: 531–536.
Lande, R. & Arnold, S.J. 1983. The measurement of selection on correlated characters. *Evolution* **37**: 1212–1226.
Lewontin, R.C. 1974. *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York.
Li, C.C. 1967. Fundamental theorem of natural selection. *Nature* **214**: 505–506.
Li, C.C. 1975. *Path Analysis*. Boxwood, Pacific Grove, CA.
Lynch, M. & Walsh, B. 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland, MA.
Mazur, B. 2008. When is one thing equal to some other thing? In: *Proof and Other Dilemmas: Mathematics and Philosophy* (B. Gold & R.A. Simons, eds), pp. 221–241. Mathematical Association of America, Washington, DC.
Michod, R.E. 1997a. Cooperation and conflict in the evolution of individuality. I. Multilevel selection of the organism. *Am. Nat.* **149**: 607–645.
Michod, R.E. 1997b. Evolution of the individual. *Am. Nat.* **150**: S5–S21.
Nowak, M.A. & Highfield, R. 2011. *SuperCooperators: Altruism, Evolution, and Why We Need Each Other to Succeed*. Free Press, New York.
Nowak, M.A., Tarnita, C.E. & Wilson, E.O. 2010. The evolution of eusociality. *Nature* **466**: 1057–1062.
Okasha, S. 2006. *Evolution and the Levels of Selection*. Oxford University Press, New York.
Page, K.M. & Nowak, M.A. 2002. Unifying evolutionary dynamics. *J. Theor. Biol.* **219**: 93–98.
Price, G.R. 1970. Selection and covariance. *Nature* **227**: 520–521.
Price, G.R. 1972a. Extension of covariance selection mathematics. *Ann. Hum. Genet.* **35**: 485–490.
Price, G.R. 1972b. Fisher's 'fundamental theorem' made clear. *Ann. Hum. Genet.* **36**: 129–140.
Price, G.R. 1995. The nature of selection. *J. Theor. Biol.* **175**: 389–396.
Queller, D.C. 1992a. A general model for kin selection. *Evolution* **46**: 376–380.
Queller, D.C. 1992b. Quantitative genetics, inclusive fitness, and group selection. *Am. Nat.* **139**: 540–558.
Rice, S.H. 2004. *Evolutionary Theory: Mathematical and Conceptual Foundations*. Sinauer Associates, Sunderland, MA.
Rice, S.H. 2008. A stochastic version of the Price equation reveals the interplay of deterministic and stochastic processes in evolution. *BMC Evol. Biol.* **8**: 262.
Robertson, A. 1966. A mathematical model of the culling process in dairy cattle. *Anim. Prod.* **8**: 95–108.
Russell, B. 1922. *Introduction to Tractatus Logico-Philosophicus by Ludwig Wittgenstein*. Harcourt, Brace and Company, New York.
Russell, B. 1958. *The ABC of Relativity*. New American Library, New York.
Schwartz, J. 2000. Death of an altruist. *Lingua Franca* **10**: 51–61.
van Veelen, M. 2005. On the use of the Price equation. *J. Theor. Biol.* **237**: 412–426.
van Veelen, M. 2011. A rule is not a rule if it changes from case to case (a reply to Marshall's comment). *J. Theor. Biol.* **270**: 189–195.
van Veelen, M., García, J., Sabelis, M.W. & Egas, M. 2010. Call for a return to rigour in models. *Nature* **466**: 661–661.
van Veelen, M., García, J., Sabelis, M. & Egas, M. 2012. Group selection and inclusive fitness are not equivalent; the Price equation vs. models and statistics. *J. Theor. Biol.* **299**: 64–80.
Wade, M.J. 1985. Soft selection, hard selection, kin selection, and group selection. *Am. Nat.* **125**: 61–73.
Weyl, H. 1983. *Symmetry*. Princeton University Press, Princeton, NJ.
Wolf, J.B., Brodie, E.D., Cheverud, J.M., Moore, A.J. & Wade, M.J. 1998. Evolutionary consequences of indirect genetic effects. *Trends Ecol. Evol.* **13**: 64–69.

Received 13 January 2012; accepted 20 February 2012

doi: 10.1111/jeb.12010

REVIEW

Natural selection. V. How to read the fundamental equations of evolutionary change in terms of information theory*

S. A. FRANK

Department of Ecology and Evolutionary Biology, University of California, Irvine, CA, USA

Keywords:

evolutionary theory;
Fisher information;
mathematical models;
population genetics.

Abstract

The equations of evolutionary change by natural selection are commonly expressed in statistical terms. Fisher's fundamental theorem emphasizes the variance in fitness. Quantitative genetics expresses selection with covariances and regressions. Population genetic equations depend on genetic variances. How can we read those statistical expressions with respect to the meaning of natural selection? One possibility is to relate the statistical expressions to the amount of information that populations accumulate by selection. However, the connection between selection and information theory has never been compelling. Here, I show the correct relations between statistical expressions for selection and information theory expressions for selection. Those relations link selection to the fundamental concepts of entropy and information in the theories of physics, statistics and communication. We can now read the equations of selection in terms of their natural meaning. Selection causes populations to accumulate information about the environment.

> There are difficulties in applying information theory in genetics. They arise principally, not in the transmission of information, but in its meaning (Maynard Smith, 2000, p. 181).

Introduction

I show that natural selection can be described by the same measure of information that provides the conceptual foundations of physics, statistics and communication. Briefly, the argument runs as follows. The classical models of selection express evolutionary rates in proportion to the variance in fitness. The variance in fitness is equivalent to a symmetric form of the Kullback–Leibler information that the population acquires about the environment through the changes in gene frequency caused by selection.

Correspondence: Steven A. Frank, Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697–2525, USA
Tel.: +1 949 824 2244; fax: +1 949 824 2181;
e-mail: safrank@uci.edu
*Part of the Topics in Natural Selection series. See Box 1.

Kullback–Leibler information is closely related to Fisher information, likelihood and Bayesian updating from statistics, as well as Shannon information and the measures of entropy that arise as the fundamental quantities of communication theory and physics. Thus, the common variances and covariances of evolutionary models are equivalent to the fundamental measures of information that arise in many different fields of study.

In Fisher's fundamental theorem of natural selection, the rate of increase in fitness caused by natural selection is equal to the genetic variance in fitness. Equivalently, the rate of increase in fitness is proportional to the amount of information that the population acquires about the environment (Frank, 2009).

In my view, information is a primary quantity with intuitive meaning in the study of selection, whereas the genetic variance just happens to be an algebraic equivalence for the measure of information. The history of evolutionary theory has it backwards, using statistical expressions of variances and covariances in place of the equivalent and more meaningful expressions of information. To read the fundamental equations of evolutionary change, one must learn to interpret the standard expressions of variances and covariances as expressions of information.

Box 1: Topics in the theory of natural selection

This article is part of a series on natural selection. Although the theory of natural selection is simple, it remains endlessly contentious and difficult to apply. My goal is to make more accessible the concepts that are so important, yet either mostly unknown or widely misunderstood. I write in a nontechnical style, showing the key equations and results rather than providing full derivations or discussions of mathematical problems. Boxes list technical issues and brief summaries of the literature.

Overview

The first section reviews the classic statistical expressions for selection. Evolutionary change caused by selection is the covariance between fitness and character value. That covariance equals the regression of character value on fitness multiplied by the variance in fitness.

The second section expresses selection in terms of the classic equations from information theory (Box 2). I show that the change in the mean logarithm of fitness is the Jeffreys information divergence. That divergence measures the accumulation of information by natural selection between the initial population and the population after it has been updated by selection. The relations between the statistical and information perspectives follow by connecting the classic statistical expressions of selection to the new information description for selection.

The third section analyses the Jeffreys divergence as the measure of information in the fundamental equations of selection. The Jeffreys divergence is the sum of two expressions for relative entropy. Relative entropy, known as the Kullback–Leibler divergence, measures the gain in information with regard to an abstract and universal notion of encoding, independently of the meaning of that information. A universal, abstract measure of information in terms of encoding allows a general theory of information to provide the foundation for the deepest concepts in communication, physics and statistics.

The fourth section concerns the meaning of information. Although encoding provides a useful measure with regard to information theory, we must also interpret the meaning of that information in terms of selection. Meaning arises by the relation of encoded information to whatever scale we use to interpret a particular problem. For selection, we interpret meaning with regard to characters. Characters may be gene frequencies or measurements made on individuals. Characters lead to a general notion of the scale for meaning with respect to the scale of encoded information.

Box 2: Information, entropy and complexity

Cover & Thomas (1991) give an excellent introduction to information theory and its applications. Jaynes (2003) is a fascinating analysis of the connections between information, entropy, probability, Bayesian analysis and statistical inference. Kullback (1959) is a broad synthesis of information theory in relation to classical statistics. Fisher's (1922, 1925) original papers on the theoretical foundations of statistics set the basis for all future work on information and statistics, with the 1925 paper showing the key role of Fisher information.

Entropy arose in the study of thermodynamics (Clausius, 1867; Boltzmann, 1872; Gibbs, 1902). Ben-Naim (2008a) gives a simple introduction. Hill (1987) provides a classical text. Information theory arose in Fisher's work and separately in the study of communication through the analyses of Hartley (1928) and Shannon (1948a, b). The underlying concepts of entropy and information are very close. Some think the concepts are identical, but controversy remains (Jaynes, 2003; Ben-Naim, 2008b).

Jeffreys (1946) divergence first appeared in an attempt to derive prior distributions for use in Bayesian analysis rather than as the sort of divergence used in this article. Kullback & Leibler (1951) and Kullback (1959) presented both the asymmetric divergence D, given in eqn 10, which is now known as the Kullback–Leibler divergence, and the symmetric form, J, given in eqn 12, which is now known as the Jeffreys divergence. They noted Jeffreys' previous usage of J in the context of Bayesian priors and then developed the importance of the divergence interpretation for statistical theory, particularly the asymmetric form, D.

I do not discuss Kolmogorov complexity in this article. However, it is an important concept that may ultimately prove as interesting for biological applications as the classic analyses of entropy and information. Kolmogorov complexity measures the information content of an object (individual) by the shortest binary computer program that fully describes the object (Cover & Thomas, 1991; Li & Vitányi, 2008). At the population level, the average Kolmogorov complexity often has a close association with the formal theories of entropy and information, but it is not exactly the same.

With respect to selection, fitness is, in essence, the match of characters to environmental challenge. That match depends on the algorithmic relation between the information content of an organism and the interpretation of that information through the development of phenotype. Development is not exactly like running a computer program encoded in the genes, but the analogy is not so far off. I suspect that, someday, Kolmogorov complexity or related measures will help to understand biochemical, developmental and evolutionary processes. A few authors have taken the first steps (Gell-Mann & Lloyd, 1996; Adami & Cerf, 2000; Adami, 2002).

The fifth section explicitly connects the abstract scale of encoded information to the meaningful scale of information in problems of selection. The analysis leads to the relation between the Jeffreys divergence, the most general expression for selection, and Fisher information as the limiting form of the Jeffreys divergence when changes in magnitude are small. Fisher information is the sensitivity of changes in abstract encoded information relative to the distance that one moves along a scale of meaning. Encoded information is equivalent to the log-likelihood ratio, which is why Fisher information provides the conceptual foundations for the theory of statistics.

The sixth section uses Fisher information to derive various elegant expressions for selection. For example, suppose that changes in the average value of a character sufficiently describe the changes caused by selection. Then, mean log fitness increases by the Fisher information in an observation about the average character value multiplied by the squared change in the average character value. This expression connects the scale of encoded information, which is mean log fitness, to the scale of meaning, which in this case is the average value of a character in the population.

The seventh section relates the parametric description of characters to a more general nonparametric expression. In the previous example, the change caused by selection was described fully by a change in a parameter, the mean. In the general case, no parametric summary statistics fully capture the change in populations. Instead, one must use the full range of different types in the population, providing a nonparametric description of the change in the distribution of frequencies caused by selection. The full nonparametric expression shows the universal applicability of the equations selection and information.

The eighth section distinguishes changes by selection from total evolutionary change. Numerous extrinsic and unpredictable forces beyond selection can change the characteristics of populations and their fit to the environment. I show the full expression for evolutionary change, placing selection in the broader evolutionary context. No general conclusion about total evolutionary change is possible, because the complete range of forces that can perturb populations remains unpredictable. However, we can express an elegant equilibrium condition. At equilibrium, the gain in information by selection must be exactly balanced by the decay in information caused by other evolutionary forces.

The discussion reviews the main argument. Classic equations for selection describe the change by statistical expressions of covariances, variances and regressions. In terms of encoded information, the change caused by selection is the Jeffreys divergence. A generalized notion of Fisher information connects encoded information to the scale of meaning. By equating the statistical description with the information description, we learn how to read the fundamental equations of selection in terms of information.

Classic equations of natural selection

Equations of natural selection are often expressed in the statistical language of population variances, covariances and regressions. In this section, I show how these statistical expressions arise from the simplest models of selection. Later sections connect these classic equations to the amount of information that a population accumulates by selection.

Textbooks on population genetics and quantitative genetics present the classic equations of selection (Crow & Kimura, 1970; Falconer & Mackay, 1996; Roff, 1997; Futuyma, 1998; Lynch & Walsh, 1998; Charlesworth & Charlesworth, 2010; Ewens, 2010). Lande developed the statistical nature of selection equations (Lande, 1979; Lande & Arnold, 1983; Frank, 1997c).

Selection

A simple model starts with n different types of individuals. The frequency of each type is q_i. Each type has w_i offspring, where w expresses fitness. In the simplest case, each type is a clone producing w_i copies of itself in each round of reproduction.

The frequency of each type after selection is

$$q_i' = q_i\left(\frac{w_i}{\bar{w}}\right), \tag{1}$$

where $\bar{w} = \sum q_i w_i$ is average fitness. The summation is over all of the n different types indexed by the i subscripts. See Box 3 for the proper interpretation of q_i'.

This equation is called a haploid model in classical population genetics, because it expresses the dynamics of different alleles at a haploid genetic locus. Recently, economists, mathematicians and game theorists have called this expression the replicator equation, because it expresses in the simplest way the dynamics of replication (Taylor & Jonker, 1978; Hofbauer & Sigmund, 1998, 2003).

It is often convenient to rewrite eqn 1 as the change in the frequency of each type, $\Delta q_i = q_i' - q_i$. Subtracting q_i from both sides of eqn 1 yields

$$\Delta q_i = q_i\left(\frac{w_i}{\bar{w}} - 1\right). \tag{2}$$

Box 3 describes a universal interpretation of these equations for selection that transcends the narrow haploid and replicator models.

Box 3: Interpretation of q' and z'

Classical population genetics and replicator equation analyses interpret q_i' in eqn 1 as the frequency of type i in the descendant population. However, selection theory in its most abstract and general form requires a set mapping interpretation, in which q_i' is the frequency of descendants derived from type i in the ancestral population. The set mapping interpretation arises from the Price equation (Price, 1972a; Frank, 1995, 1997c, 1998).

Similarly, z_i', developed in eqn 26 and mentioned earlier, is the average value of the property associated with z among the descendants derived from ancestors with index i, rather than the usual interpretation of the character value of i types in the descendant population. Here, I elaborate briefly on these interpretations of q' and z' by adapting the presentation in Frank (2012b).

Let q_i be the frequency of the ith type in the ancestral population. The index i may be used as a label for any sort of property of things in the set, such as allele, genotype, phenotype, group of individuals and so on. Let q_i' be the frequencies in the descendant population, defined as the fraction of the descendant population that is derived from members of the ancestral population that have the label i. Thus, if $i = 2$ specifies a particular phenotype, then q_2' is not the frequency of the phenotype $i = 2$ among the descendants. Rather, it is the fraction of the descendants derived from entities with the phenotype $i = 2$ in the ancestors. One can have partial assignments, such that a descendant entity derives from more than one ancestor, in which case each ancestor gets a fractional assignment of the descendant. The key is that the i indexing is always with respect to the properties of the ancestors, and descendant frequencies have to do with the fraction of descendants derived from particular ancestors.

Given this particular mapping between sets, we can specify a particular definition for fitness. Let $q_i' = q_i(w_i/\bar{w})$, where w_i is the fitness of the ith type and $\bar{w} = \sum q_i w_i$ is average fitness. Here, $w_i/\bar{w}$ is proportional to the fraction of the descendant population that derives from type i entities in the ancestors.

Usually, we are interested in how some measurement changes or evolves between sets or over time. Let the measurement for each i be z_i. The value z may be the frequency of a gene, the squared deviation of some phenotypic value in relation to the mean, the value obtained by multiplying measurements of two different phenotypes of the same entity and so on. In other words, z_i can be a measurement of any property of an entity with label, i. The average property value is $\bar{z} = \sum q_i z_i$, where this is a population average.

The value z_i' has a peculiar definition that parallels the definition for q_i'. In particular, z_i' is the average measurement of the property associated with z among the descendants derived from ancestors with index i. The population average among descendants is $\bar{z}' = \sum q_i' z_i'$.

The Price equation (eqn 26) expresses the total change in the average property value, $\Delta\bar{z} = \bar{z}' - \bar{z}$, in terms of these special definitions of set relations. This way of expressing total evolutionary change and the part of total change that can be separated out as selection is very different from the usual ways of thinking about populations and evolutionary change. The set mapping interpretation allows one to generalize equations of selection theory and total evolutionary change to a much wider array of problems than would be possible under the common interpretations of the terms. By following the set mapping approach, our evaluation of selection and information can be presented in a much simpler and more general way. Note that the classic interpretations of the haploid and replicator models are special cases of the generalized set mapping expressions.

Characters

Equation 2 describes the change in frequency. How does selection change the value of characters? Suppose that each type, i, has an associated character value, z_i. The average character value in the initial population is $\bar{z} = \sum q_i z_i$. The average character value in the descendant population is $\bar{z}' = \sum q_i' z_i'$, where z_i' is the character value in the descendants (Box 3). For now, assume that descendants have the same character value as their parents, $z_i' = z_i$. Then, $\bar{z}' = \sum q_i' z_i$, and the change in the average value of the character caused by selection is

$$\bar{z}' - \bar{z} = \Delta_s \bar{z} = \sum q_i' z_i - \sum q_i z_i = \sum (q_i' - q_i) z_i,$$

where Δ_s means the change caused by selection (Price, 1972b; Ewens, 1989; Frank & Slatkin, 1992). We may simplify this expression by using $\Delta q_i = q_i' - q_i$ for frequency changes

$$\Delta_s \bar{z} = \sum \Delta q_i z_i. \qquad (3)$$

This equation expresses the fundamental concept of selection (Frank, 2012b). Frequencies change according to differences in fitness, as given by eqn 2. Thus, eqn 3 is the change in character value caused by differences in fitness, holding constant the character values, z_i. Later, we will also include the changes in character values during transmission from parent to offspring, $\Delta z_i = z_i' - z_i$.

Variance, covariance and regression

Many of the classic equations of selection are expressed in terms of variances, covariances and regressions. I show the relation between the expression for frequency changes in eqn 3 and the common statistical expressions for selection.

Combining eqns 2 and 3 leads to

$$\Delta_s \bar{z} = \sum \Delta q_i z_i = \sum q_i \left(\frac{w_i}{\bar{w}} - 1 \right) z_i.$$

On the right-hand side, move the $\bar{w}$ term outside

$$\Delta_s\bar{z} = \sum q_i\left(\frac{w_i}{\bar{w}} - 1\right)z_i = \sum q_i(w_i - \bar{w})z_i/\bar{w}. \quad (4)$$

The definition of the population covariance allows us to rewrite this equation. Given a population of paired values (x_i, y_i), where each particular pair subscripted by i occurs at frequency q_i, and writing $\bar{x}$ as the mean value in the population of the x values, the population covariance has the general form

$$\sum q_i(x_i - \bar{x})y_i = \mathrm{Cov}(x, y).$$

Note that the right-hand expression in eqn 4 has the form of the covariance definition, so we can write

$$\Delta_s\bar{z} = \sum q_i(w_i - \bar{w})z_i/\bar{w} = \mathrm{Cov}(w, z)/\bar{w}, \quad (5)$$

following Price (1970). The standard definition of a regression coefficient of y on x is the covariance of y and x divided by the variance of x. Thus, the regression of fitness, w, on character, z, is

$$\beta_{wz} = \frac{\mathrm{Cov}(w, z)}{V_z} \quad (6)$$

where V_z denotes the variance of z. This expression implies $\mathrm{Cov}(w, z) = \beta_{wz}V_z$. We can also reverse the order of the regression, $\mathrm{Cov}(w, z) = \beta_{zw}V_w$. Thus, eqn 5 is equivalently

$$\Delta_s\bar{z} = \beta_{wz}V_z/\bar{w} = \beta_{zw}V_w/\bar{w}. \quad (7)$$

Because z can be the value of any character, we can use fitness, w, in place of z, yielding

$$\Delta_s\bar{w} = V_w/\bar{w}, \quad (8)$$

where the regression has disappeared because the regression of a variable on itself is one, thus $\beta_{ww} = 1$. This expression shows that the change in mean fitness is the variance in fitness, normalized by the initial mean value.

All of these expressions assume that character values do not change between parent and offspring, $\Delta z_i = 0$. As I mentioned, I will take up changes during transmission in a later section.

Selection expressed as change in information

This section derives a new result that connects the change in fitness caused by natural selection to the amount of information accumulated by the population. In particular, I express the change caused by selection in terms of a classical measure of information from formal information theory. Those readers unfamiliar with information theory will find some new expressions in this section, presented without explanation. The following sections explain the meaning of the expressions from information theory and the connection to natural selection. (See Boxes 4–6 for prior work on selection and information.)

Change in log fitness

Fitness captures the notion of a match between a type and the environment. We may therefore expect that fitness is, in some way, an expression of the information in the population about the environment. Those types with high fitness increase in frequency, increasing the fitness (information) contained in the population.

From eqn 1, we can write the fitness of a type, w_i, in terms of current frequencies, q_i, and updated frequencies after selection, q_i', as

$$w_i = \bar{w}\left(\frac{q_i'}{q_i}\right).$$

Fitness depends on the ratio of frequencies, q_i'/q_i. Entities that depend on ratios have a natural logarithmic scaling (Hand, 2004). Therefore, we should use the logarithmic scale when analysing fitness (Wagner, 2010). It is traditional to describe the logarithm of fitness as the Malthusian expression, $m_i = \log(w_i)$, yielding

$$m_i = \log(w_i) = \log(\bar{w}) + \log\left(\frac{q_i'}{q_i}\right).$$

Using $z \equiv m$ as our character in the selection expression of eqn 4, we have the increase in mean log fitness by natural selection as

$$\Delta_s\bar{m} = \sum \Delta q_i \log\left(\frac{q_i'}{q_i}\right). \quad (9)$$

An information measure for the change in fitness

Perhaps the most important measure of information in communication, statistics and physics is the Kullback–Leibler divergence

$$\mathcal{D}(q'\|q) = \sum q_i' \log\left(\frac{q_i'}{q_i}\right). \quad (10)$$

This divergence has directionality from the initial population, q, to the updated population after selection, q' (Box 2). Using this definition for $\mathcal{D}$ in the expression for the change in fitness given in eqn 9, we obtain

$$\Delta_s\bar{m} = \mathcal{D}(q'\|q) + \mathcal{D}(q\|q'). \quad (11)$$

This expression is the sum of Kullback–Leibler divergences taken in each direction between the initial population, q, and the updated population after selection, q'. In information theory, this sum is known as the Jeffreys divergence

Box 4: Selection and information

No one seems to have provided a full development of the relations between selection and information. In many respects, R.A. Fisher created the key concepts. However, before I start listing aspects of the problem and related citations, I cannot resist quoting from Li & Vitányi (2008, p. 96) about the difficulties of attribution. In discussing the name 'Kolmogorov complexity' for the discipline of the algorithmic analysis of complexity, they note that Solomonoff published the key idea before Kolmogorov, although Kolmogorov later discovered the idea independently and developed it more deeply and thoroughly. Ultimately, Kolmogorov got almost all the credit, perhaps because he was much more famous than Solomonoff. Li & Vitányi summarize as follows.

> Associating Kolmogorov's name with the area may be viewed as an example in the sociology of science of the Matthew effect, first noted in the Gospel according to Matthew, 25: 29–30, 'For to every one who has more will be given, and he will have in abundance; but from him who has not, even what he has will be taken away'.

Fisher (1930) discussed the relation of his fundamental theorem of natural selection to the second law of thermodynamics, a universal law about changes in entropy. However, Fisher never came around to an information perspective in this discussion and, perhaps for that reason, was restrained in his enthusiasm for the analogy. Alternatively, Fisher's restraint may have had to do with the high dimensionality of the evolutionary problem (Edwards, 2000). However, one of Fisher's great contributions in his book was his use of the average effect to reduce the dimensionality required for analysing selection. Although Fisher never developed an information analysis of selection, one must remember that the modern field of information theory only began with Shannon's work on communication (Shannon, 1948a,b). The use of Fisher information outside of statistical problems developed later.

The analogy between selection and information is obvious and has been mentioned often. However, brief mention of the analogy does not, by itself, provide any real insight about the connections between information and selection or new ways in which to understand selection.

Edwards (2000) noted that, in the continuous-time limit, the fundamental equations of selection can be expressed in terms of Fisher information. However, he concluded that the analogy between selection and Fisher information provides little insight. By contrast, Frieden *et al.* (2001) argued that selection expressed in terms of Fisher information is indeed significant. Although I believe Frieden *et al.* were on the right track, their particular analysis and presentation did not add much. Fisher information is always information about an underlying scale. Frieden *et al.* concluded that natural selection provides a measure of Fisher information about time, which I think is the wrong scale on which to interpret meaning. The present article extends the start made in Frank (2009).

$$J(q', q) = \mathcal{D}(q'\|q) + \mathcal{D}(q\|q'). \tag{12}$$

Thus, we have the simple expression for the change in mean log fitness caused by natural selection as

$$\Delta_s \bar{m} = J \tag{13}$$

where J is shorthand for $J(q', q)$. Equating this expression with eqn 7, using $m \equiv z$, we have

$$J = \beta_{wm} V_m / \bar{w} = \beta_{mw} V_w / \bar{w}, \tag{14}$$

Thus, the variance in fitness is proportional to the information divergence, J. The regression terms divided by $\bar{w}$ give the constants of proportionality that adjust for the different scales of measurement for fitness, w or m=log(w). This expression shows the relation between the information accumulated by natural selection, J, and the traditional statistical expressions of natural selection in terms of variances and regression coefficients.

The encoding of information

Before continuing to discuss the relation between selection and information, we need some additional background about the nature of information. I first describe an example in which an observation provides information. I then discuss how to quantify the amount of information. Finally, I analyse the amount of information in a comparison, which provides the basis for comparing the information in a population before and after selection.

Statistics and information

In statistical problems, the divergence, $\mathcal{D}$, measures the amount of information in an observation with respect to discriminating between two distributions (Kullback, 1959; Cover & Thomas, 1991). Suppose the true underlying probability distribution is q'. However, we do not know whether we are sampling from q' or an alternative distribution q. The different distributions may be associated with different values of a parameter, θ' and θ. The parameter may, for example, be the mean or the variance.

When we take a sample from the true underlying distribution, q', how much information do we obtain about whether the sampled distribution is q' or q? In the parametric case, how much information do we obtain about whether the parameter of the distribution from which we sampled is θ' or θ?

Box 5: Entropy, information and stochastic evolutionary models

The most interesting development of the theory arises from stochastic models of evolutionary change framed in terms of entropy and statistical mechanics. Iwasa (1988) derived a general expression for 'free fitness' by analogy with free energy and entropy. Iwasa showed the analogy between the continual increase in free fitness in evolutionary models and the second law of thermodynamics, by which entropy continually increases. He also calculated the distributions in population characteristics as they change under various stochastic models of evolutionary change.

These kinds of stochastic evolutionary models require certain assumptions in order to achieve continual increase in entropy or free fitness. There is certainly no universal law about the increase in fitness in evolution, whereas restricted notions of selection may have universal properties. I have drawn a sharp distinction between selection and evolution in my own analyses. The evolutionary literature does not always make that distinction so clearly.

de Vladar & Barton (2011a) reviewed the significant advances in the use of entropy and statistical mechanics to study evolutionary dynamics, including their own contributions to the subject (Barton & de Vladar, 2009; de Vladar & Barton, 2011b). This work on stochastic evolutionary models may eventually converge with general studies of entropy, information and dynamics. For example, there has been recent discussion about a maximum entropy production (MEP) principle for dynamics (Dewar, 2005; Kleidon, 2010; Volk & Pauluis, 2010). In the MEP theory, the most likely dynamical path is associated with the greatest production of entropy. Further, the probability distribution over dynamical paths may be a function of the relative entropy production associated with the different paths.

One may be able to use the distribution of entropy changes over paths to calculate the stochastic evolution of populations. Under some conditions, one may be able to specify the expected probability distribution over types when the population achieves certain kinds of equilibrium. However, a full understanding of MEP and its limitations has yet to be achieved. There may be some relation between dynamics analysed in terms of Fisher information (Frieden, 2004) and MEP. However, I do not understand the similarities and differences of those approaches.

For each observation, with value associated with the index i, the relative likelihood of obtaining that observation from the true distribution, q', versus the alternative distribution, q, is the ratio q'_i/q_i. The log of the likelihood ratio is $\log(q'_i/q_i)$. Because the true distribution is q', the actual probability of observing i is q'_i. Thus, averaging the log-likelihood ratio over the probability of each observed i value gives the average log-likelihood ratio, which is

$$\mathcal{D}(q'\|q) = \sum q'_i \log\left(\frac{q'_i}{q_i}\right).$$

The divergence $\mathcal{D}$ is simply the average log-likelihood ratio, which means an average of the relative weight of evidence in favour of q' as the true distribution compared with q. The greater the ratio of likelihoods, the greater the divergence between distributions and the greater the information in each observed value to discriminate between the distributions.

The scale of information

Clearly, $\mathcal{D}$ gives a measure of information provided by an observed value. But what sort of scale, or units, does that measure have? If, for example, $\mathcal{D} = 2$, then what does the value 'two' mean?

The Shannon measure of information is commonly used. That measure is related to entropy, which means randomness. The more random something is, the less information we have about it. For example, if a flipped coin comes up on either side with equal probability, we say that it is completely random. We also say that we have no information about which side is likely to come up. The Shannon measure captures this duality between increasing randomness and decreasing information or, equivalently, between decreasing randomness and increasing information.

The Shannon measure is

$$H(q) = -\sum q_i \log(q_i). \tag{15}$$

We can use any base for the logarithm. It is sometimes convenient to use base 2, in which case H is the average number of bits required to encode a message. This bit-encoding interpretation arises from the fact that

$$-\log_2(q_i) = \log_2(1/q_i)$$

expresses the number of bits required to encode a probability. For example, if q_i is 1/32, then $-\log_2(1/32) = \log_2(32) = 5$ bits. A bit is the number of digits in base two required to express a number. The number 32 in base 2 is 10000, a bit-string with 5 digits. Each digit is a bit that takes on a value of either 0 or 1.

To encode a probability 1/32 requires five bits. By contrast, to encode a probability of 1/2 requires only $\log_2(2) = 1$ bit. It takes four bits more to encode 1/32 compared with 1/2. The key idea is that a rarer event, with lower probability, q, provides greater surprise when the event actually occurs. A greater surprise means a greater distinction from what was expected, a lower ability to predict, more randomness and less information. Thus, more bits means more randomness and less information, providing a scale for measuring information in terms of bits.

Box 6: Bayesian interpretations of selection

Bayesian updating combines prior information with new information to improve prediction. The Bayesian process makes an obvious analogy with selection. The initial population encodes predictions about the fit of characters to the environment. Selection through differential fitness provides new information. The updated population combines the prior information in the initial population with the new information from selection to improve the fit of the new population to the environment. I am sure this Bayesian analogy has been noted many times. But it has never developed into a coherent framework that has contributed significantly to understanding selection.

Part of the problem is that the analogy, as currently developed, provides little more than a match of labels between the theory of selection and Bayesian theory. As Harper (2010) shows, if one begins with the replicator equation (eqn 1), then one can label the set $\{q_i\}$ as the initial (prior) population, $\{w_i/\bar{w}\}$ as the new information through differential fitness and $\{q_i'\}$ as the updated (posterior) population. Shalizi (2009) presents a similar view. The analogy provides a useful correspondence between the structure of the theories but, by itself, does not provide any truly significant insight into selection. It may be possible to develop the analogy in useful ways, a challenge that remains open.

Another Bayesian line of study analyses how individuals adjust their characters in response to information obtained directly from the environment. Those studies include learning, phenotypic plasticity, and various aspects of conditional development. By one view, learning and other processes that accumulate information follow Popper's (1972) dictum that all new knowledge must ultimately derive from trial and error, in effect, from selection.

Vast literatures discuss information theoretic and Bayesian interpretations of learning, which are beyond our scope. In an explicitly selectionist view, Fernando *et al.* (2012) analyse theories of neural development in relation to Bayesian updating – part of the wider field of developmental selection (Frank, 1996, 1997a,b). Closer to the standard evolutionary interpretation of selection, Donaldson-Matasci *et al.* (2010) provide an interesting discussion of information directly acquired from the environment in relation to fitness. Frank (1998, section 6.3) used a Bayesian analysis to combine selectively acquired information by the population as a prior state with new information acquired directly from the environment (learning).

The number of bits associated with each probability concerns only that particular probability. How should we measure the randomness and information over a set of different possible outcomes? For a distribution, q, with different probabilities q_i for each outcome, i, we must combine the randomness (bits) associated with each probability, $-\log_2(q_i)$, and the chance that the event i occurs, q_i.

In particular, the randomness associated with each event is the product of how often the event happens multiplied by the randomness of that event, $-q_i\log_2(q_i)$. The total over all events is the sum given in the definition for $H(q)$ in eqn 15, which measures the total randomness over a set of events.

To understand the notion of total randomness over a set, we can think of each i as a symbol to be communicated or an event that may occur. A message, or a set of events, has frequencies q_i. In such a set, each $-\log_2(q_i)$ is the number of bits required to encode each i, and the event i occurs with frequency q_i, so $-q_i\log_2(q_i)$ is the relative cost in terms of bits required to encode event i. If the message, or set, is highly random, it takes more bits to encode the message. High randomness corresponds to a high average level of surprise per event, which means that we have relatively little information.

Note that information is the opposite of randomness and entropy. The measurement of information can be expressed as the negative entropy, $-H$.

The information in a comparison

The problem with $-H$ as a measure of information is that, by itself, it does not give a sense of comparison or information gain. In the statistical example, we compared two distributions and the information gained to discriminate between those distributions provided by an observation. In terms of selection, we will be concerned with the information gain by a population before and after evolutionary change, requiring a comparison between the initial and updated probability distributions that describe the population before and after selection.

In a comparison, one way to measure a gain in information is by the reduction in the number of bits required to encode, or to predict, the distribution of outcomes in one population relative to another. A reduced number of bits corresponds to reduced randomness, and reduced randomness corresponds to improved prediction and more information. Thus, we can measure information gain by the reduction in the number of bits.

To make comparisons, we need an expanded definition of entropy

$$H(r,p) = -\sum r_i\log_2(p_i), \tag{16}$$

where $H(r,p)$ is the entropy in the probability distribution r when encoded by the associated probabilities p. This expression may be interpreted by thinking of the different i values as symbols in an alphabet, the r_i as

the frequency of the symbols in a message and the p_i as the frequencies used to determine the encoding of the symbols i. Then, $H(r,p)$ is the average number of bits required to encode a message r in a code based on p.

To compare populations, suppose an updated population has probabilities of types (events) q'_i, and entropy $H(q',q') = H(q')$. By contrast, the entropy of the new population, when using the encoding of the old population, q, before new information was acquired, is $H(q',q)$, which is the randomness in the new population when encoded by the old frequencies.

In the updated population, the change in information obtained from the updated encoding is the average number of bits to encode q' based on the new frequencies, $H(q',q')$, minus the average number of bits to encode q' based on the old frequencies, $H(q',q)$, which is

$$\begin{aligned} -(H(q',q') - H(q',q)) &= \sum q'_i \log_2(q'_i) - \sum q'_i \log_2(q_i) \\ &= \sum q'_i \log_2\left(\frac{q'_i}{q_i}\right) \\ &= \mathcal{D}(q'\|q), \end{aligned} \tag{17}$$

where the initial minus sign is used to express negative entropy, which is information. The term $\log_2(q'_i/q_i)$ is the number of extra bits to encode q'_i given a prior assumption that event i happens with probability q_i. The expression $\mathcal{D}$ measures the average number of extra bits needed when encoding the new population by the old frequencies rather than with the new, updated frequencies. Thus, $\mathcal{D}$ is the average gain in information in a population update when measured in terms of number of bits. A value of $\mathcal{D} = 2$ means that an efficiency gain of two bits has been achieved by the extra information provided. Alternatively, we may say that the new information enhances predictability, such that the remaining randomness, or unpredictability, has been reduced by two bits.

Selection and the meaning of information

The encoding interpretation of information is well known and widely accepted (Kullback, 1959; Cover & Thomas, 1991). By contrast, a formal interpretation of natural selection in terms of information has never been developed in a simple, clear and widely agreed manner. Here, I give my interpretation of natural selection and information.

Why *J* rather than *D* ?

To analyse the meaning of information with regard to natural selection, we must begin with the fundamental expression of selection in terms of information divergence given in eqn 13 as $\Delta_s \bar{m} = J$. That expression states that the change in mean log fitness is the Jeffreys divergence, J. Recall the definition of J from eqn 12 as

$$J(q',q) = \mathcal{D}(q'\|q) + \mathcal{D}(q\|q').$$

In most statistical and physical applications, measures of divergence and information typically use $\mathcal{D}$ (Cover & Thomas, 1991). For example, Bayesian updating can often be expressed in terms of a prior distribution, q, an updated distribution based on new data, q', and the divergence of the updated distribution from the prior, $\mathcal{D}(q'\|q)$. In the Bayesian expression, $\mathcal{D}$ describes the gain in information measured in terms of bits and interpreted with regard to the efficiency of encoding information or, equivalently, the reduced randomness and increased predictability of outcomes.

The measure $\mathcal{D}$ is asymmetric, because $\mathcal{D}(q'\|q) \neq \mathcal{D}(q\|q')$. By contrast, J is symmetric, because it is the sum of the divergence in each direction. The symmetry in the selection equation arises because, from eqn 9, we have

$$\begin{aligned} \Delta_s \bar{m} &= \sum \Delta q_i \log\left(\frac{q'_i}{q_i}\right) \\ &= \sum \Delta q_i \left[\log(q'_i) - \log(q_i)\right] \\ &= \sum \Delta q_i [\Delta \log(q_i)]. \end{aligned} \tag{18}$$

If we switch q'_i and q_i, then Δq_i changes sign and $\Delta\log(q_i)$ also changes sign. The two sign changes cancel. Thus, we obtain the same information gain when selection moves a population as $q \to q'$ or in the reverse direction as $q' \to q$.

Fitness in terms of encoded information

The information expression for fitness in eqn 18 is in terms of $\log(q'_i/q_i)$. Thus, the information gain continues to be about efficiency of encoding or, equivalently, the reduced randomness and increased predictability of outcomes. We could, for example, think of an increase in mean log fitness as an increase in the population's prediction of, or match to, the state of nature – the fit of the population to the environmental challenge.

This interpretation of fitness in terms of encoding is universal, in the sense that the particular environmental challenges and the particular meaning of the gain in fitness with respect to particular characters do not enter into the expressions. The universal expression of fitness and selection in terms of probabilities and encoding yields the match between changes in mean log fitness and changes in the classical expressions of information.

Encoding versus meaning

The great power and universality of the classic theory of information arises because it does not depend on meaning. Information is formulated strictly in terms of encoding, bits, randomness and predictability, independently of what is being encoded or predicted. Fitness

obtains the same universality, because fitness uses the same expressions of relative frequency as the classic information measures. That universality for fitness makes sense, because fitness is a general expression for the way in which populations accumulate information, independent of the characters and environmental challenges that distinguish particular cases.

Although it is certainly beneficial to have a universal expression of fitness in terms of information, we pay for that universality by the limited scope of fitness expressed only in terms of encoding. Information is about predictability, and predictability is always predictability about something. Natural selection must, in some way, be about the increased information with respect to the environmental challenges that shape success. How can we bring this particular meaning of the information about environmental challenges into the formulation of fitness?

There is perhaps no universal way to express meaning with respect to information. That may be why the encoding interpretation has been so valuable. The following sections explore two related ways in which to bring meaning into the information interpretation of fitness. The next section develops the notion of Fisher information. Later sections present the idea of a coordinate system for information and evolutionary change – a connection between the Price equation and information.

Natural selection and Fisher information

> Shannon information is not really information as such, but rather the capacity to transmit information, whereas Fisher information is truly a measure of informativeness about something specific, the value of a parameter. Shannon's refers to the medium, Fisher's to the message (Edwards, 2000, p. 6).

We have been working on the scale of encoded information. That scale depends only on probability distributions, without any explicit connection to what sort of events or meaning attach to the probabilities. Units of encoded information can be measured in terms of bits. The following extends Frank (2009).

One way to interpret meaning is to change the scale. Suppose we could relate bits of encoded information to a new scale on which we interpret meaning. To relate the change in information to the change in meaning, we could evaluate

$$\Delta\text{information} = \left(\frac{\Delta\text{information}}{\Delta\text{meaning}}\right)\Delta\text{meaning}. \qquad (19)$$

The relation is trivial when expressed in this way. However, we can see that the ratio of change in information to change in meaning provides the translation between the two scales.

To make this expression for the relations between the scales useful, we must connect each of the terms to our prior discussion of information and to a new way of describing meaning. That connection leads us to expressions of natural selection in terms of the fit of characters to the environment, rather than the efficiency of encoding information in terms of bits.

Up to this point, I have been writing q_i or q_i' for the probability of event i, whatever sort of event or characteristic i may be. The probability distribution is the set of q_i values over the range of possible characters, each possible character associated with a label i. In this formulation, one can think of the probability distributions as interpreted nonparametrically, in the sense that we work directly with the actual distribution of probabilities without reference to any underlying parameters or causes.

Now suppose we associate a set of values, θ, with each probability distribution (Amari & Nagaoka, 2000). We could think of θ as a parameter, for example the mean of the distribution. Or we could think of θ as the predictions about the environment associated with a probability distribution. The predictions might be expressed as characters. The quality of the predictions could be associated with fitness.

For now, we take θ in the general sense of some values associated with a distribution. To express the association, we expand our notation for probabilities to write $q_i|\theta$, the probability of event i given the associated value θ. An updated population may have a new associated value, θ', such as a new mean or a new prediction about the environment, so we write $q_i'|\theta'$. The change in probability is now expressed as

$$\Delta q_i|\theta = q_i'|\theta' - q_i|\theta.$$

To express the scaling of probability changes relative to changes on the new θ scale, we can divide both sides by the change on the θ scale, yielding

$$\frac{\Delta q_i|\theta}{\Delta\theta} = \frac{q_i'|\theta' - q_i|\theta}{\theta' - \theta}.$$

This expression gives us a way to match changes on the scale of meaning, θ, to changes on the scale of probability and encoded information, q.

We can now follow eqn 19 to express the change in information as the change on the scale of meaning multiplied by the change of information scaled relative to the change in meaning. To develop this expression, we must continue to match our previous work on information and selection to the new notation in relation to meaning.

The log-likelihood ratio, $\log(q_i'/q_i)$, can be written as $\log(q_i') - \log(q_i)$, which may be abbreviated as $\Delta\log(q_i)$, as in eqn 18. This difference of logarithms expresses the change in the number of bits required to encode the probabilities associated with i (as described below eqn 17). If we now express probabilities in relation to θ, as $q|\theta$, and divide by $\Delta\theta$, we obtain the change in the number of bits in relation to the change on our scale of meaning

$$\frac{\log(q_i'|\theta') - \log(q_i|\theta)}{\theta' - \theta} = \frac{\Delta\log(q_i|\theta)}{\Delta\theta}.$$

We can now put the pieces together by relating these new expressions with the expression in eqn 18 for the change in mean log fitness, yielding a form equivalent to the intuitive description in eqn 19 as

$$\Delta_s\bar{m} = \frac{J(\theta)}{\Delta\theta^2}\Delta\theta^2, \tag{20}$$

in which I write $\Delta\theta^2 = (\Delta\theta)^2$ for the square of the change in the parameter, and the term $J(\theta)$ is the Jeffreys divergence, which is now a function of the scale of meaning, θ, and is written as

$$J(\theta) = \sum(\Delta q_i|\theta)[\Delta\log(q_i|\theta)]. \tag{21}$$

These expressions simply repeat our prior derivation of $\Delta_s\bar{m} = J$, but with explicit consideration of θ.

As the changes become small, $\Delta\theta\to 0$, the Jeffreys divergence, $J(\theta)$, divided by the squared change in scale, $\Delta\theta^2$, converges to the important quantity in statistical theory known as Fisher information, $F(\theta)$, which we write as

$$\frac{J(\theta)}{\Delta\theta^2} \to F(\theta),$$

as shown in Appendix A. Thus, for small changes on the scale of meaning, $\Delta\theta\to 0$, we may write the change in average log fitness as

$$\Delta_s\bar{m} = F(\theta)\Delta\theta^2. \tag{22}$$

This derivation provides a more general way to arrive at my earlier statement that changes in mean fitness are proportional to Fisher information (Frank, 2009). Fisher information is the information in an observation about a parameter, or a set of parameters. In our case, θ represents the parameters, which is our scale of meaning.

One can also think of Fisher information as the Jeffreys divergence between populations, $J(\theta)$, relative to the squared divergence on the scale of meaning, $\Delta\theta^2$. Thus, Fisher information is the sensitivity of change in the encoded information in populations, $J(\theta)$, relative to change on the parametric scale of meaning. The greater the sensitivity, the more information in an observation with respect to the divergence between populations on the underlying parametric scale. See Appendix B for ways in which Fisher information has been used in previous models of selection.

Parametric coordinates for selection and information

The change in mean log fitness measures the amount of information that the population accumulates by selection. Because fitness describes changes in relative frequencies, fitness concerns encoding of information, which can be measured in numbers of bits.

The previous section showed how to convert from bits to an alternative scaling of information in terms of θ. We may interpret the parameters θ as a scale that has meaning with respect to the fit of the population's characteristics to the environment. This section further analyses the notion of parametric coordinates for selection and information, followed by an example.

Parametric coordinates and Fisher information

From eqn 20, the key result for the change in mean log fitness in terms of a parametric scale can be rewritten as

$$\frac{\Delta_s\bar{m}}{\Delta\theta^2} = \frac{J(\theta)}{\Delta\theta^2} \to F(\theta). \tag{23}$$

Change in mean log fitness is the amount of information gained by selection. The ratio $\Delta_s\bar{m}/\Delta\theta^2$ is the change in information per unit change in squared distance on the parametric scale. Because we consider the parametric scale as the scale of meaning, this ratio is the change in information relative to the change in squared distance on the scale of meaning (Amari & Nagaoka, 2000). The arrow on the right-hand side states that the relative change in information per unit of squared parametric distance is the Fisher information in an observation about the parameter, θ.

The interpretation of 'observation' with respect to natural selection is interesting. Each interaction of an individual with the environment leads to a realized fitness. That realized individual fitness is an observation, by the population, of the fit between certain characteristics and the environment. For a particular type, i, the average information in each observed individual fitness is $\log(q_i'/q_i) = \Delta\log(q_i|\theta)$. Thus, the ratio $\Delta\log(q_i|\theta)/\Delta\theta$ is the change, or sensitivity, of information in an observation relative to a change in θ. To get the average over all types, i, we weight this information per type by $q_i|\theta$. To analyse selection, we need the change in frequencies, or sensitivity of those changes, relative to changes in θ, which is $\Delta q_i|\theta/\Delta\theta$. Combining these terms yields $J(\theta)/\Delta\theta^2 \to F(\theta)$.

Change in the mean or variance of a character

A few examples clarify the abstract expressions for information. To keep things simple, I assume small changes so that we can use the Fisher information simplification in eqn 23. With larger changes, we could make exact calculations using $J(\theta)$ instead of Fisher information.

Change in the mean of a normal distribution under directional selection

Suppose the character values in a population, z_i, follow a normal distribution with mean, μ, and variance, v.

An observation from that population provides information about the mean of the population. It is well known that an observation from a normal population provides Fisher information about the mean of $F(\mu)=1/v$. The more variable the population, the larger v and the less information in an observation about the average value. Put another way, the precision in measurement is proportional to $1/v$. More variable populations yield less precise measurements and thus less information per observation about the average value.

We interpret natural selection as obtaining information through the observed fitnesses associated with character values. Suppose that the population retains a normal shape and a fixed variance before and after selection and changes only in its mean value. Then, the change in the mean, $\Delta\mu$, is sufficient to describe the effects of selection. From eqn 22, the increase in information by natural selection is

$$\Delta_s\bar{m} = F(\mu)\Delta\mu^2 = \frac{\Delta\mu^2}{v}.$$

This expression provides the relation between the change in information, $\Delta_s\bar{m}$, which is a universal abstract quantity about encoding, and the scaling of the character that gives meaning for this particular case, $\Delta\mu^2/v$.

Change in the variance of a normal distribution under stabilizing selection

The previous example described directional selection on the average trait value, holding the variance constant. This section considers stabilizing selection. In this case, the population begins with its centre at the optimum. Selection reduces the variance, but leaves the mean unchanged. For a normal distribution, the Fisher information in an observation about the variance, v, is $1/2v^2$. Thus,

$$\Delta_s\bar{m} = F(v)\Delta v^2 = \frac{\Delta v^2}{2v^2},$$

which is the gain in information when stabilizing selection reduces the variance of a normally distributed character.

Change in the mean of an exponential distribution

Suppose the character follows an exponential distribution before and after selection. An observation from an exponential population provides Fisher information of $1/v$ about the mean, μ. The variance of an exponential distribution is $v = \mu^2$. The change in information by selection is

$$\Delta_s\bar{m} = F(\mu)\Delta\mu^2 = \frac{\Delta\mu^2}{v},$$

which matches the case of the normal distribution. However, the variance of the exponential distribution changes with the mean. By contrast, the normal distribution has a separate parameter for the variance, which we held constant by assumption.

Change in allele frequency

Suppose $q_1 = p$ is the frequency of a particular allele and $q_0 = 1 - p$ is the frequency of the alternative allele. The distribution of allele frequencies is binomial with a single observation. The mean allelic value is $\mu = p$, and the variance is $v = p(1-p)$ The Fisher information in an observation about the mean of a binomial population is $1/v$. The change in information by selection is

$$\Delta_s\bar{m} = F(\mu)\Delta\mu^2 = \frac{\Delta\mu^2}{v}.$$

Using p for gene frequency to match the familiar notation of population genetics

$$\Delta_s\bar{m} = F(p)\Delta p^2 = \frac{\Delta p^2}{p(1-p)},$$

which holds when $\Delta\mu$=Δp is small. For larger changes, we can obtain an exact expression by using the Jeffreys divergence rather than the Fisher information, as in eqn 23.

Character coordinates and selection

The previous section assumed that the parameters, θ, summarize all differences in the frequency distributions before and after selection. We can think of θ as defining the coordinate system for evolutionary change. The reduction of frequencies to a parametric description, such as the mean of the distribution, typically requires character values to be associated with the i values. By convention, we use z_i for character values. Thus, if changes in the mean are sufficient to describe the changes in the probability distribution of characters in the population before and after selection, then $\mu = \bar{z} = \sum q_i z_i$ is a reduction of the full distribution of character values to a single parametric dimension.

Parametric character coordinates

Let us review the use of parametric coordinates before discussing nonparametric coordinates. In a parametric example, suppose that frequencies before and after selection are normally distributed, with parameters (μ,v) for the mean and the variance. Selection moves the population from the initial location, defined by the parameters (μ,v), to the location after selection, (μ',v'). The two parametric dimensions provide a complete description of change by selection. If we hold one parameter constant, such as the variance, and only allow the mean to change, then change in the single parametric dimension from μ to μ' fully describes the population before and after selection.

Parametric expressions describe the total change in information by

$$\Delta_s\bar{m} = \frac{\Delta J}{\Delta\theta^2}\Delta\theta^2 \to F(\theta)\Delta\theta^2.$$

For example, let the parameter be the mean, $\theta = \mu$. The term $J(\mu)/\Delta\mu^2 \to F(\mu)$ reduces the change in the average information per observation to the single dimension of μ. If we multiply the information per observation by the distance moved in the parametric dimension, $\Delta\mu^2$, we obtain the total change in information. Thus, the calculation for the change in information is made along the single parametric dimension of μ.

The parametric dimension of μ can be thought of as the coordinate system in which we evaluate the change by selection. Each change in position along the coordinate of μ corresponds to changes by selection, because μ is a sufficient description for the full frequency distribution of character values. In general, when we can reduce the description of frequency distributions to a sufficient set of parameters, θ, those parameters form the coordinates in which we evaluate the changes by selection.

Nonparametric character coordinates

We can think of our fundamental expression for selection

$$\Delta_s\bar{z} = \sum \Delta q_i z_i$$

as a nonparametric expression. Each term includes the actual frequencies in the population. The calculation is made over the full dimensionality of the frequency distribution.

The character values, $\{z_i\} = z_1, z_2, \ldots$, form a nonparametric coordinate system. For the population frequencies, $\{q_i\}$, the point $\{q_i z_i\}$ locates the population before selection and the point $\{q_i' z_i\}$ locates the population after selection. The movement of the population caused by selection is given by $\{\Delta q_i z_i\}$.

The expression for the total change in information caused by selection is

$$\Delta_s\bar{m} = J = \sum \Delta q_i \Delta \log(q_i) = \sum \Delta q_i \log\left(\frac{q_i'}{q_i}\right).$$

Each frequency change, Δq_i, associates with the character $z_i = \Delta\log(q_i)$, the change in information for the ith type. This is a nonparametric expression, because the calculation is made over the full frequency distribution.

Character coordinates and information

The character values provide the coordinates of meaning in an analysis of selection. We can derive the relations between information and the coordinates of meaning by using the results of eqns 7 and 8. From those equations, we obtain the relation between the change given the coordinates of meaning, $\Delta_s\bar{z}$, and the change given the coordinates of information, $\Delta_s\bar{m}$, as

$$\Delta_s\bar{z} = \left(\frac{\beta_{zw}}{\beta_{mw}}\right)\Delta_s\bar{m}. \tag{24}$$

The term β_{zw} is the regression coefficient of the character values, z, on the fitnesses, w. The term β_{mw} is the regression coefficient of the log fitnesses, m, on the fitnesses, w. These regressions provide an exact expression for changing the coordinates from information, $\Delta_s\bar{m}$, to characters, $\Delta_s\bar{z}$. When the magnitudes of the changes are small, $w \to m+1$, thus

$$\Delta_s\bar{z} \to \beta_{zm}\Delta_s\bar{m}. \tag{25}$$

To repeat, it is important to recognize a regression coefficient as an exact expression for the change in scale associated with a change in coordinates. The regression is sufficient when evaluating the consequences for a change in coordinates with respect to a change in mean value.

The underlying values, z_i, may themselves be nonlinear functions of other values (Frank, 2012b). For example, z_i could be the product of different character values measured on each individual, or the square of some underlying character. What matters is that we average over the z_i values to get $\Delta_s\bar{z}$.

Character coordinates and total evolutionary change

The previous analyses have focused on the selection part of total evolutionary change. I defined selection as the change caused by frequency differences

$$\Delta_s\bar{z} = \sum \Delta q_i z_i.$$

The subscript s emphasizes that this expression is the partial change caused by selection (Price, 1972b; Ewens, 1989; Frank & Slatkin, 1992).

Total change in characters

The partial change arises by holding constant the character values, such that $\Delta z_i = z_i' - z_i = 0$. This assumption fixes the coordinates, z_i, and evaluates the meaning of changing frequencies in the context of that fixed set of coordinates.

If the coordinates that give meaning also change, $\Delta z_i \neq 0$, then we must account for that change in coordinates with respect to the total evolutionary change. In particular, the total change is the sum of the change, Δ_s, caused by selection through varying frequencies, q, holding constant the coordinates, z, plus the change in coordinates, Δ_c, holding constant the new frequencies in the updated population, q'. We write the total change as

$$\Delta\bar{z} = \Delta_s\bar{z} + \Delta_c\bar{z} = \sum \Delta q_i z_i + \sum q_i' \Delta z_i. \qquad (26)$$

This expression is a form of the Price equation. I devoted the prior article to a full discussion of this equation (Frank, 2012b). Here, I focus only on those aspects that concern information. In particular, I emphasize the interpretation of z as a coordinate system that gives meaning to the information basis of natural selection.

Total change in information

The total evolutionary change in eqn 26 can be used to evaluate information. Let $z = m$, where the log fitness, m, provides a measure of the information accumulated by a population. Thus,

$$\Delta\bar{m} = \Delta_s\bar{m} + \Delta_c\bar{m}. \qquad (27)$$

From eqn 13, the selection component of change is $\Delta_s\bar{m} = J$. In general, no simplified reduction or particular interpretation is possible for the change in coordinates, $\Delta_c\bar{m}$. That change in coordinates arises from any environmental or extrinsic factors that may change, altering the fit of the characters to the environment. The changes in the frequencies themselves can be an 'environmental' change that alters fitnesses (Price, 1972b; Ewens, 1989; Frank & Slatkin, 1992). Thus, no general expression for total evolutionary change in fitness is possible other than

$$\Delta\bar{m} = J + \Delta_c\bar{m}.$$

One can, of course, analyse particular models such as mutation–selection balance. Mutation decays information through changes in fitness that are, on average, negative, causing a loss of information through the term $\Delta_c\bar{m} = \sum q_i' \Delta m_i$. The particular loss of information through $\Delta_c\bar{m}$ depends on the specific assumptions. By contrast, the gain in information through selection is always $\Delta_s\bar{m} = J$.

Equilibrium balance between information gain and loss

Many processes lead to an equilibrium balance between gain of information by selection and decay of information by an opposing force (Frank, 2012a). Mutation–selection balance is one example. Frequency-dependent selection is another, in which the gain in information by selection is balanced by the decay of information (fitness) caused by frequency changes. For example, in the evolution of sex ratios, making more daughters may be favoured by selection. But as the number of daughters increases by selection, the advantage of making extra daughters decays.

Although we cannot, in general, specify the change in the coordinate term, $\Delta_c\bar{m}$, we can express the equilibrium condition, $\Delta\bar{m} = 0$. Under a balance between information gain by selection and information decay by change in coordinates,

$$J = -\Delta_c\bar{m}.$$

It is sometimes possible to analyse particular problems by using that universal expression for the balance of forces (Frank & Slatkin, 1990; Frank, 1995).

Evolution of the coordinate system

In the previous sections, I have fixed the particular dimensions that define the coordinate system. Although the coordinates may change, Δz_i, each dimension i remained. From a broader perspective, the evolution of the various dimensions in the coordinate system itself is perhaps among the most interesting evolutionary problems. One aspect concerns the origin of new characters (West-Eberhard, 2003). More generally, one may consider the evolution of the optimal set of characters with respect to the capture of information.

There is an interesting literature in engineering about optimal design of sensors with respect to capturing information. That literature sometimes uses Fisher information as the optimality criterion with respect to design (Borguet & Léonard, 2008). Application of that design perspective with regard to information may provide insight into biological problems. For example, multiple cellular receptors may respond to the same sort of information, such as the concentration of a hormone. But those receptors may be tuned differently with regard to sensitivity to signals. A related idea concerns the common trade-off between informativeness and simplicity in classification (Kemp & Regier, 2012).

A second aspect of coordinates concerns the parametric reduction of the full nonparametric distribution of characters. Reducing the full distribution to the mean is an extreme reduction and probably not justified in general. However, there often may be some suitable reduction of dimensionality to a sufficient set of parameters with respect to the acquisition of information (Carter *et al.*, 2009; Goh *et al.*, 2011). That sufficient set defines the coordinates of information and meaning followed by an evolving population. It may be that an improved parametric representation of information in the environment by a set of characters enhances fitness. Thus, it may be the parametric representation itself that is under the strongest selection or, at least, a particularly interesting form of selection.

Discussion

The fundamental equations of selection are often written in the statistical terms of variances, covariances and regressions. I have argued that one obtains a deeper understanding of selection if one learns to read the fundamental equations in terms of information. Here, I review my argument by listing the key steps derived

in previous sections. I start with the classic statistical equations of selection. I then show the connection of those statistical expressions of selection to expressions for the information that populations accumulate about the fit of characters to the environment.

Statistical expressions of selection

To understand where the classic statistical expressions of selection come from and what they mean, let us start with the basic equation for evolutionary change by natural selection

$$\Delta_s \bar{z} = \sum \Delta q_i z_i$$

given in eqn 3. Here, $\Delta_s \bar{z}$ is the change caused by selection in the average value of a character, $\bar{z}$. This expression applies generally to selection of any value. For example, z could be gene frequency, leading to population genetics expressions, or z could be a quantitative trait such as weight, or z could be a nonlinear function of several characters. The Δq_i terms are the changes caused by selection in the frequency of the ith character value, z_i. Total selection is the total change in frequencies, with each change caused by selection, Δq_i, weighted by its associated character value, z_i.

I showed that one can rewrite the association between the change caused by selection and the character value as

$$\sum \Delta q_i z_i = \mathrm{Cov}(w, z)/\bar{w}, \tag{28}$$

a form known as the Price equation and also related to Robertson's secondary theorem of natural selection (Frank, 2012b). This form provides the foundation for quantitative genetics theory and also arises in standard models of population genetics. The definition of covariance allows us to rewrite the covariance as the product of a regression coefficient and a variance term

$$\Delta_s \bar{z} = \mathrm{Cov}(w, z)/\bar{w} = \beta_{zw} V_w/\bar{w}, \tag{29}$$

where β_{zw} is the regression of character value, z, on fitness, w, and V_w is the variance in fitness. These sorts of regression and variance terms arise repeatedly in the fundamental equations of selection.

One can easily understand why selection depends on an association between fitness, w, and character value, z. Those character values associated with higher fitness will increase, whereas those character values associated with lower fitness will decrease. But why should the expression for selection be exactly the covariance, or the regression multiplied by the variance, which capture only the linear component of association? The reason is that $\Delta_s \bar{z}$ describes selection by a change in average values. To calculate a change in the average, we need only the linear component of association between character and fitness.

These statistical expressions of selection in terms of covariances, variances and regressions have been very useful throughout the history of evolutionary theory. However, these expressions give no sense of what selection means. To say that selection is the covariance of fitness and character value is simply to express an algebraic relation. That algebraic relation is very useful, but it does not give a sense of what selection is actually doing with regard to adaptation or how selection relates to processes in other fields of study. The statistical expressions do not tell us how to read the fundamental equations of selection with regard to the meaning of the underlying process.

Selection in terms of information

In this article, I argued that selection causes populations to accumulate information about the fit of characters to the environment. I gave a precise definition of 'information'. That definition of information with respect to selection matches exactly the classic usage of information and entropy from the fundamental theories of physics, statistics and communication. By showing the exact relations between selection and information, I tied the theory of natural selection to the broader conceptual framing of problems at the foundation of many key scientific disciplines.

I will not repeat the whole argument here. Instead, I list a few steps to emphasize the essential points. To understand the information associated with selection and fitness, we must analyse fitness on a logarithmic scale

$$m_i = \log(w_i) = \log(\bar{w}) + \log\left(\frac{q_i'}{q_i}\right).$$

The logarithmic scale compares relative magnitudes. We need relative magnitudes because there is no meaning in the number of babies or the number of copies produced with regard to whether a type, i, is increasing or decreasing in the population. We need to know the relative success. The logarithmic scale is the natural scale of relative magnitudes.

Using log fitness, m, as the character value of interest in eqn 28, we obtain

$$\Delta_s \bar{m} = \sum \Delta q_i m_i = \sum \Delta q_i \log\left(\frac{q_i'}{q_i}\right).$$

We recognize the fundamental expression for the change in information given by the Kullback–Leibler divergence, or relative entropy, as

$$\mathcal{D}(q' \| q) = \sum q_i' \log\left(\frac{q_i'}{q_i}\right).$$

Using this definition for change in information, $\mathcal{D}$, we can express the change in mean log fitness caused by selection as

$$\Delta_s \bar{m} = \mathcal{D}(q' \| q) + \mathcal{D}(q \| q').$$

This sum of the changes in information in each direction is known as the Jeffreys divergence, J. Thus, we can write the fundamental expression for the accumulation in information by natural selection as

$$\Delta_s \bar{m} = J.$$

Because z in eqn 29 is just a placeholder for any character, we can use m in place of z in that equation, yielding

$$\Delta_s \bar{m} = \beta_{mw} V_w / \bar{w}.$$

Thus, the information accumulated by natural selection is equivalently expressed in terms of the regression coefficient and variance

$$J = \beta_{mw} V_w / \bar{w}. \tag{30}$$

The value of J is the gain in information. The variance in fitness, V_w, is therefore a measure of the separation between the initial population and the population after selection, when the separation between populations is expressed on a scale of information. The regression divided by the mean fitness, $\beta_{mw}/\bar{m}$, is a scaling factor that translates the measure of information in V_w to the scale of log fitness, m. That scaling change is required because log fitness is the proper measure of information in expressions of selection.

Equation 30 shows the equivalence between the expression of information gain and the expression of it in terms of statistical quantities. There is nothing in the mathematics to favour either an information interpretation or a statistical interpretation.

I have argued that, when reading the fundamental equations of selection for meaning, we should prefer the information interpretation. The information perspective makes sense intuitively. Selection is the process by which populations accumulate information about the environment.

Acknowledgments

My research is supported by National Science Foundation grant EF-0822399.

References

Adami, C. 2002. What is complexity? *BioEssays* **24**: 1085–1094.

Adami, C. & Cerf, N.J. 2000. Physical complexity of symbolic sequences. *Physica D* **137**: 62–69.

Amari, S. & Cichocki, A. 2010. Information geometry of divergence functions. *Bull. Polish Acad. Sci. Tech. Sci.* **58**: 183–195.

Amari, S. & Nagaoka, H. 2000. *Methods of Information Geometry*. Oxford University Press, New York.

Barton, N.H. & de Vladar, H.P. 2009. Statistical mechanics and the evolution of polygenic quantitative traits. *Genetics* **181**: 997–1011.

Ben-Naim, A. 2008a. *Entropy Demystified: The Second Law Reduced to Plain Common Sense*. World Scientific, Singapore.

Ben-Naim, A. 2008b. *A Farewell to Entropy: Statistical Thermodynamics Based on Information*. World Scientific, Singapore.

Boltzmann, L. 1872. Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen. *Wien. Akad. Sitz* **66**: 275–370.

Borguet, S. & Léonard, O. 2008. The Fisher information matrix as a relevant tool for sensor selection in engine health monitoring. *Int. J. Rotat. Machine.* **2008**: 784749.

Carter, K.M., Raich, R., Finn, W. & Hero, A. 2009. FINE: Fisher information nonparametric embedding. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**: 2093–2098.

Charlesworth, B. & Charlesworth, D. 2010. *Elements of Evolutionary Genetics*. Roberts & Company, Greenwood Village, CO.

Cichocki, A. & Amari, S. 2010. Families of alpha- beta- and gamma-divergences: Flexible and robust measures of similarities. *Entropy* **12**: 1532–1568.

Cichocki, A., Cruces, S. & Amari, S. 2011. Generalized alpha-beta divergences and their application to robust nonnegative matrix factorization. *Entropy* **13**: 134–170.

Clausius, R. 1867. *Mechanical Theory of Heat*. John Van Voorst, London.

Cover, T.M. & Thomas, J.A. 1991. *Elements of Information Theory*. Wiley, New York.

Crow, J.F. & Kimura, M. 1970. *An Introduction to Population Genetics Theory*. Burgess, Minneapolis, MN.

Dewar, R.C. 2005. Maximum entropy production and the fluctuation theorem. *J. Phys. A Math. Gen.* **38**: L371-L381.

Donaldson-Matasci, M.C., Bergstrom, C.T. & Lachmann, M. 2010. The fitness value of information. *Oikos* **119**: 219–230.

Edwards, A.W.F. 2000. Fisher information and the fundamental theorem of natural selection. *Istit. Lomb. (Rend. Sc.) B* **134**: 3–17.

Ewens, W.J. 1989. An interpretation and proof of the fundamental theorem of natural selection. *Theor. Popul. Biol.* **36**: 167–180.

Ewens, W.J. 1992. An optimizing principle of natural selection in evolutionary population genetics. *Theor. Popul. Biol.* **42**: 333–346.

Ewens, W.J. 2010. *Mathematical Population Genetics: I. Theoretical Introduction*, 2nd edn. Springer-Verlag, New York.

Falconer, D.S. & Mackay, T.F.C. 1996. *Introduction to Quantitative Genetics*, 4th edn. Longman, Essex, England.

Fernando, C., Szathmáry, E. & Husbands, P. 2012. Selectionist and evolutionary approaches to brain function: a critical appraisal. *Front. Comput. Neurosci.* **6**: 24.

Fisher, R.A. 1922. On the mathematical foundations of theoretical statistics. *Philos. Trans. R. Soc. Lond. A* **222**: 309–368.

Fisher, R.A. 1925. Theory of statistical estimation. *Math. Proc. Cambridge Phil. Soc.* **22**: 700–725.

Fisher, R.A. 1930. *The Genetical Theory of Natural Selection*. Clarendon, Oxford.

Frank, S.A. 1995. George Price's contributions to evolutionary genetics. *J. Theor. Biol.* **175**: 373-388.

Frank, S.A. 1996. The design of natural and artificial adaptive systems. In: *Adaptation* (M.R. Rose & G.V. Lauder, eds), pp. 451–505. Academic Press, San Diego, CA.

Frank, S.A. 1997a. The design of adaptive systems: optimal parameters for variation and selection in learning and development. *J. Theor. Biol.* **184**: 31–39.

Frank, S.A. 1997b. Developmental selection and self-organization. *BioSystems* **40**: 237-243.

Frank, S.A. 1997c. The Price equation, Fisher's fundamental theorem, kin selection, and causal analysis. *Evolution* **51**: 1712–1729.

Frank, S.A. 1998. *Foundations of Social Evolution*. Princeton University Press, Princeton, NJ.
Frank, S.A. 2009. Natural selection maximizes Fisher information. *J. Evol. Biol.* **22**: 231–244.
Frank, S.A. 2012a. Natural selection. III. Selection versus transmission and the levels of selection. *J. Evol. Biol.* **25**: 227–243.
Frank, S.A. 2012b. Natural selection. IV. The Price equation. *J. Evol. Biol.* **25**: 1002–1019.
Frank, S.A. & Slatkin, M. 1990. The distribution of allelic effects under mutation and selection. *Genet. Res.* **55**: 111–117.
Frank, S.A. & Slatkin, M. 1992. Fisher's fundamental theorem of natural selection. *Trends Ecol. Evol.* **7**: 92–95.
Frieden, B.R. 2004. *Science From Fisher Information: A Unification*. Cambridge University Press, Cambridge, UK.
Frieden, B.R., Plastino, A. & Soffer, B.H. 2001. Population genetics from an information perspective. *J. Theor. Biol.* **208**: 49–64.
Futuyma, D.J. 1998. *Evolutionary Biology*, 3rd edn. Sinauer Associates, Sunderland, MA.
Gell-Mann, M. & Lloyd, S. 1996. Information measures, effective complexity, and total information. *Complexity* **2**: 44–52.
Gibbs, J.W. 1902. *Elementary Principles in Statistical Mechanics*. Scribner, New York.
Goh, A., Lenglet, C., Thompson, P.M. & Vidal, R. 2011. A nonparametric Riemannian framework for processing high angular resolution diffusion images and its applications to odf-based morphometry. *NeuroImage* **56**: 1181–1201.
Hand, D.J. 2004. *Measurement Theory and Practice*. Arnold, London.
Harper, M. 2010. The replicator equation as an inference dynamic. arXiv:0911.1763v3.
Hartley, R.V.L. 1928. Transmission of information. *Bell Syst. Tech. J.* **7**: 535–563.
Hill, T.L. 1987. *An Introduction to Statistical Thermodynamics*. Dover, New York.
Hofbauer, J. & Sigmund, K. 1998. *Evolutionary Games and Population Dynamics*. Cambridge University Press, Cambridge, UK.
Hofbauer, J. & Sigmund, K. 2003. Evolutionary game dynamics. *Bull. Am. Math. Soc.* **40**: 479–519.
Iwasa, Y. 1988. Free fitness that always increases in evolution. *J. Theor. Biol.* **135**: 265–281.
Jaynes, E.T. 2003. *Probability Theory: The Logic of Science*. Cambridge University Press, New York.
Jeffreys, H. 1946. An invariant form for the prior probability in estimation problems. *Proc. R. Soc. London A* **186**: 453–461.
Kemp, C. & Regier, T. 2012. Kinship categories across languages reflect general communicative principles. *Science* **336**: 1049–1054.
Kimura, M. 1958. On the change of population fitness by natural selection. *Heredity* **12**: 145–167.
Kleidon, A. 2010. A basic introduction to the thermodynamics of the earth system far from equilibrium and maximum entropy production. *Philos. Trans. R. Soc. Lond. B* **365**: 1303–1315.
Kullback, S. 1959. *Information Theory and Statistics*. Wiley, New York.
Kullback, S. & Leibler, R.A. 1951. On information and sufficiency. *Ann. Math. Stat.* **22**: 79–86.
Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution* **33**: 402–416.
Lande, R. & Arnold, S.J. 1983. The measurement of selection on correlated characters. *Evolution* **37**: 1212–1226.
Li, M. & Vitányi, P. 2008. *An Introduction to Kolmogorov Complexity and its Applications*. Springer-Verlag, New York.
Lynch, M. & Walsh, B. 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland, MA.
Maynard Smith, J. 2000. The concept of information in biology. *Philos. Sci.* **67**: 177–194.
Popper, K.R. 1972. *Objective Knowledge: An Evolutionary Approach*. Oxford, Oxford, UK.
Price, G.R. 1970. Selection and covariance. *Nature* **227**: 520–521.
Price, G.R. 1972a. Extension of covariance selection mathematics. *Ann. Hum. Genet.* **35**: 485–490.
Price, G.R. 1972b. Fisher's 'fundamental theorem' made clear. *Ann. Hum. Genet.* **36**: 129–140.
Roff, D.A. 1997. *Evolutionary Quantitative Genetics*. Chapman and Hall, New York.
Shalizi, C.R. 2009. Dynamics of Bayesian updating with dependent data and misspecified models. *Electron. J. Stat.* **3**: 1039–1074.
Shannon, C.E. 1948a. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**: 379–423.
Shannon, C.E. 1948b. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**: 623–656.
Taylor, P.D. & Jonker, L.B. 1978. Evolutionary stable strategies and game dynamics. *Math. Biosci.* **40**: 145–156.
de Vladar, H.P. & Barton, N.H. 2011a. The contribution of statistical physics to evolutionary biology. *Trends Ecol. Evol.* **26**: 424–432.
de Vladar, H.P. & Barton, N.H. 2011b. The statistical mechanics of a polygenic character under stabilizing selection, mutation and drift. *J. R. Soc. Interface* **8**: 720–739.
Volk, T. & Pauluis, O. 2010. It is not the entropy you produce, rather, how you produce it. *Philos. Trans. R. Soc. Lond. B* **365**: 1317–1322.
Wagner, G.P. 2010. The measurement theory of fitness. *Evolution* **64**: 1358–1376.
West-Eberhard, M.J. 2003. *Developmental Plasticity and Evolution*. Oxford University Press, New York.

Appendix A: Fisher information as the limiting form of the Jeffreys divergence

A large family of divergence measures converges to Fisher information in the limit of small changes (Amari & Nagaoka, 2000; Amari & Cichocki, 2010; Cichocki & Amari, 2010; Cichocki *et al.*, 2011). In this appendix, I show that the limit of the Jeffreys divergence is the Fisher information multiplied by a scaling factor for parametric distance.

I also show that the chi-square divergence becomes the Fisher information metric in the limit of small changes. The different forms of divergence can be confusing if one does not realize that all of the different divergence measures in the Fisher family are equivalent in the limit, but differ when changes are not small.

My main point is that the Jeffreys divergence holds the unique position as the only correct divergence measure for models of selection. It is the only measure that is correct both for large changes and, in the limit, for

small changes. As far as I know, my derivation in this article of the Jeffreys divergence in relation to selection has not been shown previously. The clear relation of the Jeffreys divergence to changes in information is essential to make the proper connection between selection and information.

Limiting form of Jeffreys divergence

I show $J(\theta) \to F(\theta)\Delta\theta^2$ as the distance in the parametric coordinates $\Delta\theta^2 \to 0$. Notationally, $\Delta\theta^2 \equiv (\Delta\theta)^2$. Using the standard differential notation for small differences, we write $\Delta\theta^2 \to \mathrm{d}\theta^2$. Thus, I show $J(\theta) \to F(\theta)\mathrm{d}\theta^2$.

I use the vector θ as parametric coordinates for probability distributions, following standard analysis in information geometry (Amari & Nagaoka, 2000). For simplicity, I usually treat the parametric vector as a single dimension. The extension to multiple dimensions is standard.

The Jeffreys divergence in parametric form, from eqn 21, is

$$J(\theta) = \sum (\Delta q_i|\theta)[\Delta \log(q_i|\theta)].$$

As the changes become small, $\Delta q_i|\theta = q_i'|\theta' - q_i|\theta \to 0$ and $\Delta\theta = \theta' - \theta \to 0$, we write

$$\begin{aligned}\Delta q_i|\theta &\to \mathrm{d}q_i|\theta \\ &= \left(\frac{\mathrm{d}q_i|\theta}{\mathrm{d}\theta}\right)\mathrm{d}\theta \\ &= \dot{q}_i \mathrm{d}\theta,\end{aligned}$$

where $\dot{q}_i$ is the derivative of $q_i|\theta$ with respect to θ. Next,

$$\begin{aligned}\Delta \log(q_i|\theta) &\to \mathrm{d}\log(q_i|\theta) \\ &= \left(\frac{\mathrm{d}\log(q_i|\theta)}{\mathrm{d}\theta}\right)\mathrm{d}\theta \\ &= \left(\frac{\dot{q}_i}{q_i}\right)\mathrm{d}\theta,\end{aligned}$$

where, to make the notation more concise, I use $q_i \equiv q_i|\theta$. Thus,

$$J(\theta) \to \sum \left(\frac{\dot{q}_i^2}{q_i}\right)\mathrm{d}\theta^2.$$

Below, I show that $\sum \dot{q}_i^2/q_i$ is Fisher information, $F(\theta)$. Thus, $J(\theta) \to F(\theta)\mathrm{d}\theta^2$.

Pearson's chi-square divergence

We have from the previous expression

$$J(\theta) \to \sum \left(\frac{\dot{q}_i^2}{q_i}\right)\mathrm{d}\theta^2 = \sum \frac{\mathrm{d}q_i^2}{q_i}. \tag{31}$$

Pearson's chi-square divergence, or chi-square test statistic, is usually described as follows. Given an expected probability distribution, $\{q_i\}$, and an observed probability distribution, $\{q_i'\}$, the chi-square statistic is the sum of observed minus expected squared over expected. Writing the observed minus expected squared as $\Delta q_i^2 = (q_i' - q_i)^2$, we have

$$\chi^2(\theta) = \sum \frac{\Delta q_i^2}{q_i}.$$

As the changes become small,

$$\chi^2(\theta) \to \sum \frac{\mathrm{d}q_i^2}{q_i} = \sum \left(\frac{\dot{q}_i^2}{q_i}\right)\mathrm{d}\theta^2,$$

demonstrating that the Jeffreys and chi-square divergences have the same limiting form. The next section shows that the limiting form is related to the Fisher information metric.

When changes are large, only the Jeffreys divergence gives the correct expression for changes by selection in mean log fitness, $\Delta_s\bar{m}$. The chi-square divergence is the change in mean fitness on a linear scale

$$\Delta_s\bar{w} = \sum \Delta q_i w_i = \sum \frac{\Delta q_i^2}{q_i}.$$

As I discussed in the text, the correct scale for analysing the changes in fitness is logarithmic, because fitness is a relative measure, and logarithmic scaling is the correct scale for relative measures (Wagner, 2010). In addition, the relations between selection and information are only clear on the logarithmic scale, because it is only on that scale that one can see the connections to the classic theories of entropy and information. In the limit of small changes, the logarithmic scale becomes linear, and thus, $\Delta_s\bar{m} \to \Delta_s\bar{w}$.

Alternative expressions for Fisher information

One can think of Fisher information as the change in a probability distribution with respect to a change in a parameter that specifies the distribution. The more rapidly a distribution changes with respect to a parameter, the more information each observation provides about the value of the parameter. For example, if the distribution changes very slowly, then small differences in the distribution of observed values may translate into big differences in parameter values. Thus, approximately similar distributions of observations map to widely different parameter values, so each observation provides relatively little information about the parameter. If, by contrast, the distribution changes rapidly with respect to a parameter, then the distribution of observations is very different for small changes in the parameter, and each observation provides much information about the likely value of the parameter.

Mathematically, Fisher information is the negative value of the expected curvature of the log-likelihood function

$$F(\theta) = -\sum q_i \left(\frac{\mathrm{d}^2 \log(q_i|\theta)}{\mathrm{d}\theta^2} \right).$$

Doing the differentiation, and noting (Amari & Nagaoka, 2000) that

$$\sum \frac{\mathrm{d}^2 q_i|\theta}{\mathrm{d}\theta^2} = \frac{\mathrm{d}}{\mathrm{d}\theta} \sum \frac{\mathrm{d}q_i|\theta}{\mathrm{d}\theta} = 0,$$

because the sum of changes in frequencies must be zero over a distribution, we obtain

$$F(\theta) = \sum \frac{\dot{q}_i^2}{q_i}.$$

A large number of different divergence measures converge to Fisher information in the limit. Thus, knowing only that the limiting form of a divergence is Fisher information only weakly constrains the associated form of divergence. For example, from the expression above for the chi-square divergence

$$\chi^2(\theta) \to \sum \frac{\mathrm{d}q_i^2}{q_i} = \sum \left(\frac{\dot{q}_i^2}{q_i} \right) \mathrm{d}\theta^2,$$

it might be tempting, in a particular application in which Fisher information arises, to think of the chi-square divergence as somehow the natural measure of change, because the chi-square form for large changes most closely resembles the limiting Fisher information form for small changes. In the case of selection, that conclusion would not be correct. The Jeffreys divergence is in fact the natural measure of change, because the logarithmic scale is the natural scale for changes in fitness and for changes in information.

Appendix B: Historical aspects

Kimura (1958) noted that the change in fitness in certain models of selection is

$$\Delta_s \bar{m} = \sum \frac{\dot{q}_i^2}{q_i}. \tag{32}$$

Kimura used the standard notion of change with respect to time in his study of continuous dynamics with respect to small changes. Thus, the parameter is $\theta \equiv t$ for time, and $\dot{q} = \mathrm{d}q/\mathrm{d}t$.

Ewens (1992) and Edwards (2000) provide comprehensive syntheses of the literature on the various uses of Kimura's expression, $\sum \dot{q}_i^2/q_i$. The main use concerned information geometry expressions of selection dynamics on a Riemannian manifold. Neither Ewens nor Edwards found that discussion of information geometry particularly useful. Edwards did note that the Kimura's expression is in fact just an expression for Fisher information. But Edwards did not think that association was useful.

I agree with the criticisms by Ewens and Edwards within the context of how the literature had been framed. From Kimura (1958) through the various developments in the literature, the emphasis had always been on dynamics with respect to time. I agree with Edwards that one cannot say anything very interesting about the temporal dynamics of evolutionary change from the simple expression in eqn 32 for selection. That expression is the partial change caused by selection (Price, 1972b; Ewens, 1989; Frank & Slatkin, 1992), not the total evolutionary change. The partial change gives a clear sense of what selection is doing at any moment, but provides no insight by itself about evolutionary dynamics.

My presentation in this article is also based on Fisher information and, more generally, on the Jeffreys divergence. Two aspects of my presentation go beyond the past work and, in my view, provide a compelling case for framing our understanding of selection in these terms.

First, I connected selection to information theory through the general result $\Delta_s \bar{m} = J$, the Jeffreys divergence. This result does not depend on the limit of small changes, but instead is a general description of the nature of selection. This result establishes the proper measure for the amount of information accumulated by selection.

Second, I related the change in information to various underlying parametric and nonparametric scales. Those scales provide the meaning with respect to the abstract scale for encoded information that forms the basis for classical information theory. As Edwards (2000) emphasized, Fisher information is information about meaning with respect to underlying parameters (Frank, 2009). Earlier work implicitly used time as the parameter, which is not a meaningful way of expressing the accumulation of information. One does not think of selection as providing information about time. In addition to making the parametric basis for selection and information explicit, my use of the Jeffreys divergence clarified the relation of selection to classical information theory.

Finally, I achieved greater generality than past work by respecting the fundamental distinction between selection and evolution. Past work often tried to make general statements about evolutionary dynamics, which is not possible. It is possible to make strong and completely general statements about the partial change caused by selection. Such statements clarify the relations between selection and information. One can achieve that depth and generality only by working within the fundamental limitations imposed by the distinction between selection and total evolutionary change.

I mentioned that Ewens (1992) and Edwards (2000) concluded that past work based on the Kimura's result

did not contribute significantly to understanding selection. Ewens (1992) did develop his own extension to that theory, in which he showed an optimization principle in relation to Fisher's fundamental theorem. Frank (2009) developed a similar idea but with a different approach that emphasized information and the Fisher information metric. Those studies derive from a partitioning of the causes of fitness, which is the topic of a future article in this series.

Received 6 July 2012; revised 15 September 2012; accepted 20 September 2012

JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.12066

REVIEW

Natural selection. VI. Partitioning the information in fitness and characters by path analysis*

S. A. FRANK

Department of Ecology and Evolutionary Biology, University of California, Irvine, CA, USA

Keywords:

information theory;
kin selection;
mathematical models;
path analysis;
population genetics.

Abstract

Three steps aid in the analysis of selection. First, describe phenotypes by their component causes. Components include genes, maternal effects, symbionts and any other predictors of phenotype that are of interest. Second, describe fitness by its component causes, such as an individual's phenotype, its neighbours' phenotypes, resource availability and so on. Third, put the predictors of phenotype and fitness into an exact equation for evolutionary change, providing a complete expression of selection and other evolutionary processes. The complete expression separates the distinct causal roles of the various hypothesized components of phenotypes and fitness. Traditionally, those components are given by the covariance, variance and regression terms of evolutionary models. I show how to interpret those statistical expressions with respect to information theory. The resulting interpretation allows one to read the fundamental equations of selection and evolution as sentences that express how various causes lead to the accumulation of information by selection and the decay of information by other evolutionary processes. The interpretation in terms of information leads to a deeper understanding of selection and heritability, and a clearer sense of how to formulate causal hypotheses about evolutionary process. Kin selection appears as a particular type of causal analysis that partitions social effects into meaningful components.

> The path method … is not so much concerned with prediction as [it is with] the proposal of a *plausible interpretation* of the relationships between the variables. In other words, path analysis is concerned with erecting a causal structure compatible with the observed data (Li, 1975, p. 3).

Introduction

Populations accumulate information by natural selection. The amount of information may be expressed by classical information theory (Frank, 2012b). That purely informational expression describes phenotypes and fitness abstractly, without consideration of the explicit causes that determine phenotypic traits and their association with fitness. Here, I partition phenotypes and fitness into their component causes.

For phenotypes, we must track the influence of genes, symbionts, maternal effects and other potential causes. The components of phenotype lead to explicit models of character expression and heritability. For fitness, we must track how different characters and external forces combine to determine success. An individual's fitness may, for example, depend on a combination of its own phenotype and the phenotypes of its neighbours.

I put those explicit causal components of phenotype and fitness into the fundamental expressions of selection and evolutionary change. I recover an expanded concept of heritability, a precise understanding of Fisher's fundamental theorem and a general form of the equations of selection for multiple characters. With those tools, the following article clarifies kin selection and other social processes (Frank, 2013).

Correspondence: Steven A. Frank, Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697–2525, USA. Tel.: +1 949 824 2244; fax: +1 949 824 2181;
e-mail: safrank@uci.edu

*Part of the Topics in Natural Selection series. See Box 1.

Box1: Topics in the theory of natural selection

This article is part of a series on natural selection. Although the theory of natural selection is simple, it remains endlessly contentious and difficult to apply. My goal is to make more accessible the concepts that are so important, yet either mostly unknown or widely misunderstood. I write in a non-technical style, showing the key equations and results rather than providing full derivations or discussions of mathematical problems. Boxes list technical issues and brief summaries of the literature.

I presented much of this material in Frank (1997b, 1998). Here, I pursue four goals. First, I express the key partitions of phenotypes and fitness with respect to my new information theory interpretation of selection (Frank, 2012b). Second, the information expressions translate the traditional regression and variance terms of selection into more meaningful descriptions of cause and consequence. Third, the partitions of phenotype and fitness provide the basis for replacing outdated concepts of kin selection with a solid conceptual foundation (in Frank, 2013). Fourth, I emphasize simplicity, presenting the mathematical material at the most basic level consistent with the concepts. The original publications contain more detail (Frank, 1997b, 1998).

Mathematically, little is required beyond simple forms of statistical regression and the location of points in coordinate systems. Although I use only basic mathematics, the article is nonetheless challenging. I cover a wide array of problems at a very general level, with emphasis on the connections between seemingly different topics. That sustained abstraction and synthesis provide both significant rewards and demanding challenges.

It may seem that the basic problems of selection and kin interactions were solved long ago. Why do we need to revisit those topics? In fact, our understanding of natural selection and kin selection has continued to advance over the past few decades. Those advances have developed while the old formulations have remained. The core of the subject has become cluttered with incompatible expressions from different eras, derived in different contexts. One can no longer go forward without first resetting the foundations.

Selection

I briefly review the general equations for selection and evolution. Recent articles in this series provide full details (Frank, 2012a, b).

The Price equation

Consider an initial population. Let $\bar{z}$ be the average in the population of some value (phenotype). A second population has average value $\bar{z}'$. Total change between the populations is $\Delta\bar{z} = \bar{z}' - \bar{z}$. Split the total change into two components

$$\Delta\bar{z} = \Delta_s\bar{z} + \Delta_c\bar{z}. \quad (1)$$

The first term, Δ_s, is the part of the total change caused by selection. The second term, Δ_c, is the remaining part of total change by all other causes.

To evaluate these terms, we write the average value as $\bar{z} = \sum q_i z_i$. The index i divides the population in any way that we choose. We may use i to label by different individuals, by different groups, by genotype or by any other partition of the population. The frequency of a type i in the population is q_i. The phenotype associated with i is z_i. The average value in the second population is $\bar{z}' = \sum q_i' z_i'$.

We define selection as changes in frequency, holding constant phenotype

$$\Delta_s\bar{z} = \sum q_i' z_i - \sum q_i z_i.$$

Here, the populations differ in their frequencies, $\Delta q_i = q_i' - q_i$, but we have held the phenotype values constant at z_i in both populations. Using Δq_i for frequency change, we write

$$\Delta_s\bar{z} = \sum \Delta q_i z_i. \quad (2)$$

To obtain the total change, we need the changes in phenotype holding constant the frequencies

$$\Delta_c\bar{z} = \sum q_i' z_i' - \sum q_i' z_i.$$

Here, the populations differ in their phenotype, $\Delta z_i = z_i' - z_i$, but we have fixed the frequency at q_i'. We use the final frequencies in the second population, q', because they provide the proper reference for final phenotype after change (Box 2). Using Δz_i for phenotypic changes, we write

$$\Delta_c\bar{z} = \sum q_i' \Delta z_i.$$

The total change from eqn 1 can now be written (Box 2) as a form of the Price equation

$$\Delta\bar{z} = \sum \Delta q_i z_i + \sum q_i' \Delta z_i. \quad (3)$$

Box 2: Price equation: difference of a product

The Price equation simply expands a difference into multiple terms. Consider, for example, the difference of the product of x and y, which we write as $\Delta(xy) = x'y' - xy$. We can expand the difference of the product as

$$\Delta(xy) = (x + \Delta x)(y + \Delta y) - xy$$

which yields

$$\Delta(xy) = (\Delta x)y + x(\Delta y) + \Delta x \Delta y.$$

This expression shows that the difference of a product is the difference of the first term holding the second term constant, plus the difference of the second term holding the first term constant, plus the product of the two differences.

We can simplify the difference expansion by combining a pair of terms on the right-hand side. Noting that $x' = x + \Delta x$, we can combine the last two terms into one, yielding

$$\Delta(xy) = (\Delta x)y + x'(\Delta y).$$

The derivation of the Price equation follows the rule for the difference of a product

$$\begin{aligned}\Delta\bar{z} &= \Delta \sum q_i z_i \\ &= \sum \Delta(q_i z_i) \\ &= \sum (\Delta q_i) z_i + \sum q_i'(\Delta z_i).\end{aligned}$$

The value of the Price equation arises from identifying $\sum(\Delta q_i)z_i$ as the part of total change caused by selection. Selection acts on phenotype at a fixed point in time, so it makes sense to consider selection as the partial difference in frequency holding phenotype constant. When we use log fitness for the phenotype, $m \equiv z$, we get an exact correspondence between the selection term and the increase in information expressed by classical information theory (eqn 8). That correspondence supports interpreting $\sum(\Delta q_i)z_i$ as selection.

Classical expressions of covariance, regression and variance

The definition of fitness is

$$q_i' = q_i \frac{w_i}{\bar{w}}, \tag{4}$$

where w_i is the fitness of type i, and $\bar{w}$ is average fitness. The change in frequency is

$$\Delta q_i = q_i\left(\frac{w_i}{\bar{w}} - 1\right).$$

Thus, the change caused by selection can be written as a covariance between fitness and phenotype

$$\sum \Delta q_i z_i = \sum q_i\left(\frac{w_i}{\bar{w}} - 1\right)z_i = \mathrm{Cov}(w, z)/\bar{w}. \tag{5}$$

We can rewrite a covariance as a product of a regression coefficient and a variance term

$$\Delta_s\bar{z} = \mathrm{Cov}(w, z)/\bar{w} = \beta_{zw}V_w/\bar{w}, \tag{6}$$

where β_{zw} is the regression of phenotype, z, on fitness, w and V_w is the variance in fitness. Selection equations are often expressed with these covariance, regression and variance terms. Classical population genetics expressions for change in gene frequency also have this form, in which we let $\bar{z} = p$ be the frequency of a gene in a population.

Information

Frank (2012b) showed that selection can be expressed in terms of information theory. I briefly review the key points in this section.

Fitness and the gain in encoded information

Fitness, w, describes relative changes in frequency. Logarithms provide the natural scaling for relative changes. Using the expression for fitness in eqn 4, we write log fitness as

$$m_i = \log(w_i) = \log(\bar{w}) + \log\left(\frac{q_i'}{q_i}\right).$$

Using $z \equiv m$ in the expression for selection (eqn 2), we have

$$\Delta_s\bar{m} = \sum \Delta q_i m_i = \sum \Delta q_i \log\left(\frac{q_i'}{q_i}\right).$$

The classic information theory expression for the change in encoded information between two populations with frequencies q' and q is

$$\mathcal{D}(q'\|q) = \sum q_i' \log\left(\frac{q_i'}{q_i}\right). \tag{7}$$

With that definition, we have

$$\Delta_s\bar{m} = \mathcal{D}(q'\|q) + \mathcal{D}(q\|q'),$$

in which the right-hand side is known as the Jeffreys information divergence, J. Thus, we can write the fundamental expression for the accumulation of information by natural selection as

$$\Delta_s\bar{m} = J. \tag{8}$$

Because z in eqn 6 is just a placeholder for any character, we can use m in place of z in that equation, yielding

$$\Delta_s \bar{m} = \beta_{mw} V_w / \bar{w}.$$

Thus, the information accumulated by natural selection, J, is equivalently expressed in terms of the regression coefficient and the variance,

$$J = \beta_{mw} V_w / \bar{w}. \tag{9}$$

Variance, regression and information

The variance in fitness, V_w, is proportional to the information gain by natural selection, J (eqn 9). It is easy to understand why selection may be expressed in terms of information. Selection is, in essence, a process by which populations gain information about the environment. But, why should the variance arise as an alternative description of selection?

The usual view is that selection acts on differences within the population. The greater the differences, the larger the variance and the greater the opportunity for selection. But, why exactly is the variance the correct measure of differences within the population, rather than some other measure of variation?

Consider the definition of fitness (eqn 4) given earlier

$$\frac{w_i}{\bar{w}} = \frac{q_i'}{q_i},$$

in which the relative fitness is the ratio of frequencies between the new and old population. Relative fitness is, in essence, a measure of the separation between the new population and the old population, a comparison of q' vs. q. Because the frequencies in each population must add to one, each separation between a pair q_i' and q_i must be balanced by opposite separations in other pairs.

Thus, the variation in the q_i'/q_i ratios measures the total separation of the new population from the old population. In particular, the variance in those ratios – the variance in fitness – is like a distance between the new population and the old population. That distance-like measure has units in terms of the information gain (Frank, 2012b). The variance in fitness expresses an informational distance, the amount of information gained by selection.

Information gain is measured on the logarithmic scale of frequency changes (eqn 7). The regression coefficient, β_{mw}, transforms fitness from the linear scale, w, to the log scale, m, yielding the key expression given earlier for the change in log fitness (information) caused by selection

$$\Delta_s \bar{m} = J = \beta_{mw} V_w / \bar{w}.$$

It is common to think of a regression coefficient as a linear prediction estimated from data. That interpretation misleads with regard to understanding the fundamental equations of selection. Instead, the regression coefficient describes the consequence for the change in average value when transforming from one scale to another scale (Boxes 3 and 4). The proper way to read β_{mw} is a change in scale from w to m when evaluating the averages $\bar{w}$ and $\bar{m}$.

Phenotype as a change in the scaling of information

Selection causes populations to accumulate information. The measure of information is related to log fitness. In the analysis of selection, we often focus on phenotypes rather than fitness. Here, I show that, with respect to selection, one can think of the phenotypic scale simply as an alternative scale on which to measure information.

Begin with the expression given earlier for the change in log fitness

$$\Delta_s \bar{m} = \beta_{mw} V_w / \bar{w}.$$

The regression coefficient, β_{mw}, changes scale from fitness, w, to log fitness, m. If we divide by β_{mw}, we obtain

$$\frac{\Delta_s \bar{m}}{\beta_{mw}} = V_w / \bar{w}.$$

The factor $1/\beta_{mw}$ reverses the scale change, transforming from the logarithmic scale, m, to the linear scale, w.

The change in phenotype from eqn 6 can be written as

$$\Delta_s \bar{z} = \beta_{zw} V_w / \bar{w}.$$

The regression β_{zw} changes scale from fitness, w, to phenotype, z, and $1/\beta_{zw}$ reverses the direction of the change in scale. Thus

$$\frac{\Delta_s \bar{z}}{\beta_{zw}} = V_w / \bar{w} = \frac{\Delta_s \bar{m}}{\beta_{mw}}.$$

Because the information accumulated by natural selection is $\Delta_s \bar{m} = J$, we have

$$\Delta_s \bar{z} = \left(\frac{\beta_{zw}}{\beta_{mw}} \right) J.$$

This expression describes the change in phenotype by selection in relation to the information gain, J, rescaled by the transformation from the scale of information, m, to the scale of phenotype, z. We may describe the scaling between the gain in information, J, and change in phenotype caused by selection, $\Delta_s \bar{z}$, as

$$\alpha_z = \frac{\beta_{zw}}{\beta_{mw}}. \tag{10}$$

Thus we can write the relation between the change in phenotype and the gain in information as

$$\Delta_s \bar{z} = \alpha_z J. \tag{11}$$

Box 3: Regression

Simple regression is based on the equation for a line

$$z = a + \beta y,$$

in which z is the outcome of interest, y is a variable that is used to predict z, the term β is the slope of the line relating z to y and a is the intercept, which is the value of z when $y = 0$. The simple regression model is usually written as

$$z_i = a + \beta_{zy} y_i + \delta_i,$$

in which the i subscripts denote values associated with different observations, and δ_i is the residual as described below. In some applications, it is convenient to make the intercept a disappear, which we achieve by $y_i = x_i - a/\beta_{zy}$, which gives

$$z_i = \beta_{zx} x_i + \delta_i.$$

This expression is equivalent to the previous one. The only change is that x differs from y by a constant value. The second expression uses β_{zx} in place of β_{zy}. Those terms have the same value, but I use the term with x to emphasize that the relation is now between z and x. In any regression model, we can make a similar substitution in which we change y by a constant factor to get an x value that makes the intercept disappear.

From the perspective of regression analysis, $\beta_{zx}x$ provides a prediction of z given x. The difference between the actual value and the predicted value is the residual (error), $\delta_i = z_i - \beta_{zx}x$. Two changes in notation provide a cleaner expression. Write the regression coefficient as $b = \beta_{zx}$, and drop the i subscript, yielding

$$z = bx + \delta,$$

where the variables implicitly range over i.

Regression has a natural asymmetry. In prediction, the value of z is the predicted value given the predictor, x. In a causal interpretation, in the sense of path analysis (Box 5), the effect z depends on the cause, x. One must keep this asymmetry in mind to interpret regression equations correctly. Proper notation helps. We may write

$$z \,|\, x = bx + \delta,$$

which emphasizes that the outcome, z, depends on the given fixed value of x. We read $z \mid x$ as 'z given x'. If we take the average of both sides

$$\mathrm{E}(z \,|\, x) = bx,$$

where $\mathrm{E}(z \mid x)$ is the expectation of z given x, in which 'expectation' means the average value. On the right side, δ disappears because the regression coefficient, b, is chosen so that the average value of the residual is zero, $\bar{\delta} = 0$.

Causes of phenotype

This section partitions the causes of phenotype into components. The next section connects the causes of phenotype to the capture and transmission of information. The following section partitions fitness into components, dividing the gain in information by selection into different causes. Boxes 3–6 provide background on regression. Box 7 provides citations to the literature.

Overview

Heritability describes the expected similarity in phenotype between different individuals (Falconer & Mackay, 1996). For example, we may define the predictors of phenotype as the set of alleles in an individual, and the heritability as the part of similarity between ancestors and descendants ascribed to those alleles. Because sex and recombination break up particular combinations of alleles, adding up the effects of each individual allelic predictor often provides a good estimate of the similarity between different relatives caused by genetics.

Alternatively, we may expand the set of predictors to include certain nonlinear combinations of alleles. For example, we may have a predictor for the presence of allele A, another for the presence of allele B, and a third for the presence of both alleles. Certain expanded predictor sets may give a more accurate description of similarity between closely related ancestor–descendant pairs that are likely to share the allelic combinations, but may give a less accurate description when the allelic pairs tend to be broken up during transmission.

Here, I am primarily interested in the information that a population accumulates by selection, and how different processes may reduce or alter the transmission of accumulated information. My expressions include the classic genetic measures as special cases. But, I do not emphasize the connection to traditional genetics – the genetic interpretations are discussed in every basic textbook of genetics (Falconer & Mackay, 1996). Instead, I focus on general equations for selection and the transmission of information. In my expressions, any predictors can be used including, but not limited to, all of the traditional genetic forms.

Why bother with such abstractions? Because many extensions to basic genetic theory have been developed to cope with nongenetic effects or to analyse selection independently of genetics (Lynch & Walsh, 1998). The

Box 4: Change in scale

In the regression model (Box 3) with subscripts used explicitly for labelling types

$$\mathrm{E}(z_i \,|\, x_i) = bx_i.$$

If we consider subscripts for two different types, k and i, we can write $\mathrm{E}(z_k \,|\, x_k) = bx_k$ and $\mathrm{E}(z_i \,|\, x_i) = bx_i$. Subtracting these two equations from each other gives

$$\mathrm{E}(z_k - z_i \,|\, x_k - x_i) = b(x_k - x_i).$$

Using Δ to denote a change between the k and i values

$$\mathrm{E}(\Delta z \,|\, \Delta x) = b(\Delta x),$$

which we can write equivalently as

$$b = \beta_{zx} = \frac{\mathrm{E}(\Delta z \,|\, \Delta x)}{\Delta x},$$

which we read as: 'the regression of z on x is the expected change in z for a given change in x divided by the change in x'. From this expression, we see that a regression coefficient is the expected change in scale for one variable in relation to another variable. One can also think of the regression coefficient as a sort of generalization of differentiation. For situations in which we can consider z and x as continuous variables with an underlying functional relationship, $z(x)$, it will often be the case that as the changes become small, $\Delta z \to 0$ and $\Delta x \to 0$ with x confined to a small range of values, then the regression coefficient approaches the derivative, $\beta_{zx} \to \mathrm{d}z/\mathrm{d}x$.

Finally, the variables x and δ are uncorrelated, so that $\mathrm{Cov}(x,\delta) = 0$. Regression uses all of the available information in x about z. Thus, any left over deviations, δ, cannot contain information about z, which is reflected in the lack of correlation between those variables.

When we have multiple predictors, or causes, $x_j = x_1, x_2, \ldots, x_n$, then the regression equation is

$$z = \sum_j b_j x_j + \delta,$$

where each b_j is the partial regression of z on x_j, holding constant the other predictor values. Suppose, for example, that we have two predictors, x_1 and x_2. For notational convenience, let $x \equiv x_1$ and $y \equiv x_2$, so that the regression equation is

$$z = b_x x + b_y y + \delta.$$

If, as above, we take the difference between two x values, holding y constant, we obtain

$$b_x = \beta_{zx \cdot y} = \frac{\mathrm{E}(\Delta z | \Delta x, y)}{\Delta x},$$

which we read as: 'the regression of z on x, holding y constant, is the expected change in z for a given change in x and a fixed value of y, divided by the change in x'. This expression gives the expected change in scale between z and x for a given value of y. If z, x and y are continuous variables with an underlying functional relationship, $z(x,y)$, then for small changes confined to a small range of predictor values for x and y, it will often be the case that the regression approaches the partial derivative $\beta_{zx \cdot y} \to \partial z/\partial x$.

These properties of regression follow from least squares. The squared distance between predicted and observed values is the sum of squares, $\sum \delta_i^2$. Minimizing that distance gives the least value for the sum of squares – the least squares. All properties here follow from that minimization. Further aspects of regression depend on other assumptions. For example, many tests of statistical significance assume that the residuals have a normal distribution. Certain interpretations require that the observations be linearly related to the predictors. I do not use those further aspects and therefore do not require any assumptions about linearity or the distribution of observations and residuals.

literature tends to deal with each particular problem as a novel challenge that requires special theory. For example, maternal effects, kin selection, cultural evolution and institutional evolution in economics all have their distinct literatures and ways of framing problems. Yet all of those problems are just examples of a general theory of selection and transmission. In any particular application, the key is to express the causes of phenotypes (characteristics) and the causes of fitness (success) by a model, or hypothesis, of how various predictors combine to determine outcome. A general theory expressed in terms of any choice of predictors defines the unifying conceptual framework (Frank, 1997b, 1998).

Fisher's average effect

We can separate phenotype into components by

$$z_i = \sum_j b_j x_{ij} + \delta_i.$$

Each type, i, has n different associated x_i values, $x_{i1}, x_{i2}, \ldots, x_{in}$. From the perspective of multiple regression, the x's are predictors, or independent variables, with respect to the phenotype, z. Each b_j is a partial regression coefficient of z on x_j. Roughly speaking, a partial regression coefficient, b_j, describes the average change in phenotype, z, for a change in the associated predictor variable, x_j.

Box 5: Causes and predictors

> Since path analysis depends on structure, and structure in turn depends on the cause-and-effect relationship among the variables, we shall first say a few words about the way these terms will be used … There are a number of formal definitions as to what constitutes a cause and what an effect. For instance, one may think that a cause must be doing something to lead to something else (effect). While this is clearly one type of cause-and-effect relationship, we shall not limit ourselves to that type only. Nor shall we enter into philosophical discussions about the nature of cause-and-effect. We shall simply use the words 'cause' and 'effect' as statistical terms similar to independent and dependent variables, or [predictor variables and response variables] (Li, 1975, p. 3).

I analyse causes of phenotypes and causes of fitness. Here, I briefly comment on the word 'cause'. The above quote and the epigraph come from Li's book on *Path Analysis*. Li's point concerns the distinction between three levels of analysis. First, true causality describes the relations between actual forces and actual effects. Whether such things can ever be studied or known directly remains a philosophical problem beyond our scope.

Second, at the other extreme, multiple regression analysis from classical statistical theory concerns only correlations and variances. The standard theory explicitly disavows causal interpretation – correlation is not causation. Regression arises by minimizing the distance between predicted outcomes and actual outcomes – an attempt at optimal prediction. One thinks of the variables used to predict outcome simply as predictors that, in the past, would have helped one to make a better guess about what actually happened. The predictors may have direct effects themselves or be correlated with some other unseen causal factor. However, those notions of direct and unseen cause are irrelevant to the method.

Third, path analysis takes an intermediate approach. One chooses the predictors for a model as a hypothesis about cause. Rather than aim for optimal prediction, one aims for a set of variables that consistently describe the observed patterns of variation. The quality of the causal interpretation is primarily evaluated by the consistency of the hypothesized pathways in capturing the observed variance in outcome. Consistency roughly means relative stability in the magnitude of a pathway's effect under different circumstances. Although that interpretation potentially offers some insight into cause and effect, the analytical method remains multiple regression. One simply emphasizes the quality of a model as a potential causal interpretation rather than as an attempt at optimal prediction.

Consider a model in which we use genes as predictors of phenotype. In a breeding programme to improve yield, we want to predict offspring phenotype to make the best choice of breeding design. Causality is irrelevant, we aim only for a good outcome. By contrast, in a theoretical analysis of adaptation by natural selection, we want to understand the causal processes. How do the genes that affect phenotype combine to determine morphology or behaviour? How does selection influence the underlying genes and the resulting phenotypic design in relation to performance? We are after an understanding of the process. The quality of prediction will, of course, be the primary way to interpret the causal model. But a good prediction arising from the wrong underlying causal model is what we most want to avoid. Prediction becomes a method for evaluation rather than the goal.

This article analyses natural selection in relation to causal interpretations. For that reason, I think of my models of multiple regression as models of path analysis. In a different context, the same models could be thought of strictly as analyses of regression and prediction.

We often focus on the general relation of a phenotype, z, to its components, x_j, rather than on the particular phenotype, z_i, of a particular type, i, in relation to its particular components, x_{ij}. Thus, we may express the general relation between a phenotype and its components as

$$z = \sum_j b_j x_j + \delta,$$

in which one understands that the particular values of z, x_j and δ vary for the different types, i, whereas the average effect of a predictor, b_j, is a property of the population.

The regression expression applies to any predictors, x_j. We could use temperature, neighbours' behaviour, another phenotype, epistatic interactions given as the product of allelic values, symbiont characters or an individual's own genes. Fisher first presented this regression for phenotype in terms of alleles. Suppose each x_j is the presence or absence of an allelic type. Then each b_j describes the average contribution to phenotype for adding or subtracting the associated allelic type, and b_j is called the average effect (Fisher, 1930; Crow & Kimura, 1970; Falconer & Mackay, 1996).

Predicted phenotype is

$$g = \sum_j b_j x_j. \tag{12}$$

In genetic contexts, g is often called the breeding value (Falconer & Mackay, 1996). Using g, we can partition phenotype into a predicted component and a residual component

$$z = g + \delta, \tag{13}$$

where $\delta = z - g$ is the difference between the actual value and the predicted value. If we take the average of both sides, we get $\bar{z} = \bar{g}$, because $\bar{\delta} = 0$.

Box 6: Nonlinearity

Regression and path analysis are sometimes thought to be limited to linear and additive effects. However, that is misleading. Consider $z = bx + \delta$. Here, b is the linear relation between x and z. However, it may be that $x = y^2$, in which the true underlying cause is y. Thus, we are actually regressing on a nonlinear function of a causal variable, y. Or, it may be that we start with $z = b_1x_1 + b_2x_2 + b_3x_3 + \delta$. This appears to be an additive model. However, the underlying cause may be $x_1 = y_1$, and $x_2 = y_2$ and $x_3 = y_1 * y_2$. Thus, our model expresses nonlinearity and nonadditivity in the causes, y.

In general, any nonlinear relation can be expressed by an additive sum of terms, in which the individual terms may be nonlinear. Thus, regression can fully account for any nonlinearity by an additive sum of terms. In practice, limitations arise because we may not know the correct nonlinear relation, and so cannot express the proper sum of nonlinear terms. However, that is not a limitation of regression, but rather a limitation that arises from our ignorance. Another method of analysis does not solve the problem of our ignorance. The point is that one must distinguish limitations arising from method from limitations arising from ignorance. Confusing those different limitations is a common mistake.

The components of heritability

The part of phenotype not transmitted

Typically, we only follow the transmission of the predictors. For example, we may follow transmission of genes plus any other variables we choose. Those effects that we include explicitly end up as part of the predicted phenotype, g, and as candidates for the transmitted phenotype. All effects on phenotype not explicitly included as predictors end up in the residual, δ. The split between the predicted phenotype and the residual is arbitrary. If we add a new predictor, any additional effect of that predictor moves from the residual, δ, to the predicted phenotype, g. Usually, we wish to give the best description of the causes of phenotype that we can. Thus, our choice of predictors defines our hypothesis about the causes of phenotype, in the sense of path analysis discussed in Box 5.

The part of phenotype associated with the particular set of predictors, g, defines one component of heritability. Aspects of phenotype not associated with the particular predictors in our model appear as a nontransmitted component of phenotype, δ, reducing the similarity of phenotype between ancestors and descendants associated with the predictors.

Change in transmitted components of phenotype

A second component of heritability arises from the stability of the effects associated with the predictors. If a predictor has effect bx in the original population and effect $b'x'$ in the second population, then the transmission of that predictor is associated with a change in

Box 7: Brief history of evolutionary partitions

Fisher (1918, 1930) partitioned phenotype into its various genetic causes. Quantitative genetics extended the partitioning of phenotype by genetic and nongenetic causes (Falconer & Mackay, 1996; Lynch & Walsh, 1998). Models of cultural evolution use culturally transmissible attributes as predictors of phenotype (Dawkins, 1976; Cavalli-Sforza & Feldman, 1981; Boyd & Richerson, 1985).

Quantitative genetic models may also consider partitions of fitness into component causes. Recent work on partitions of fitness was stimulated by Lande & Arnold (1983). Many subsequent studies expanded that approach, including various explicit descriptions based on path analysis (Heisler & Damuth, 1987; Crespi & Bookstein, 1989; Crespi, 1990; Kingsolver & Schemske, 1991; Scheiner *et al.*, 2000). I unified the different lines of study on partitions of phenotype and partitions of fitness (Frank, 1997b, 1998), motivated initially by Queller's quantitative genetic models of kin selection (Queller, 1992a, b).

In the text, I mentioned that $rB - C > 0$ can sometimes be interpreted in terms of group selection. For example, if neighbours' phenotype, y, is an average character value in a local group, then r can be defined as the regression of individual character value on group character value. That group regression can be considered in a path analysis model, which is roughly the way in which Heisler & Damuth (1987) analysed group selection. In their article, they emphasized 'contextual analysis' similarly to the way in which I have emphasized 'path analysis'. Frank (1995b) and Taylor & Frank (1996) also calculated r by regressing group value on individual value in several models, following a long tradition that blurred the mathematical distinction between kin and group selection (Hamilton, 1975; Frank, 1986).

Some of the multivariate analyses of fitness attempt to predict evolutionary dynamics, and therefore must make explicit assumptions about the distribution of phenotypes and the nature of heritability. I do not discuss dynamics; my models do not require any of those extra assumptions.

phenotype $\Delta(bx) = b'x' - bx$. Box 2 shows that we can express this change as

$$\Delta(bx) = (\Delta b)x + b'(\Delta x).$$

Summing over the j different predictors and using the definition of g from eqn 12 yields

$$\Delta g = \sum(\Delta b_j)x_j + \sum b'_j(\Delta x_j). \tag{14}$$

On the right side, the first term describes the change in the predicted value of a type that arises from the changes in the average effects of the predictors, Δb_j, holding constant the predictor values, x_j. For example, the average effect of an allele on phenotype may be frequency dependent. Thus, the average effect will change over time as the frequency of the allele changes in the population. The second term describes the change in the transmitted predictor values, Δx_j, evaluated in the context of the average effects from the second population, b'_j. For example, an allele may mutate into another form, thus weighting the average effect by a different amount.

The smaller the Δb and Δx values, the less the phenotype changes with respect to the transmitted predictors, and the higher the heritability associated with those predictors. Equivalently, the more stable the predictors and their average effects, the greater the fidelity at which those particular predictors transmit the information accumulated by selection to the new population.

The change in the predictors, Δx, includes mutation as well as any other process that alters predictor values (Frank, 1995a, 1997b, 1998; Price, 1995). For example, predictors in a descendant may derive from multiple ancestors. We can think of the mixing of predictors by considering the change in predictor values when derived from different sources. In some cases, we may wish to alter the assignment of descendants to ancestors. For example, a behaviour may influence the frequency of nondescendant types. To associate the behavioural phenotype with the change in frequency, we could assign those nondescendants to the ancestral behaviour responsible for their presence (Hamilton, 1970). In general, we can make such assignments in any way that we choose. The key is that assigning different descendants to an ancestor may alter the change in predictor values between a descendant and its assigned ancestor. Such changes may alter the fidelity at which information is transmitted (Frank, 1998). I will take up that topic in the next article (Frank, 2013).

The part transmitted and the change during transmission

The full, exact expression from eqn 3 for the total evolutionary change is

$$\Delta\bar{z} = \sum \Delta q_i z_i + \sum q'_i \Delta z_i.$$

We can partition phenotype as $z = g+\delta$, the split between the part explained by the predictors of phenotype, g, and the part that is not explained by the set of predictors in our model for phenotype, δ. From eqn 13, $\Delta\bar{z} = \Delta\bar{g}$ because $\bar{\delta} = 0$, thus

$$\Delta\bar{z} = \Delta\bar{g} = \sum \Delta q_i g_i + \sum q'_i \Delta g_i.$$

With $g_i = z_i - \delta_i$, we get

$$\Delta\bar{z} = \sum \Delta q_i z_i - \sum \Delta q_i \delta_i + \sum q'_i \Delta g_i. \tag{15}$$

We can express each of these terms with a particular notation that emphasizes its interpretation

$$\Delta\bar{z} = \Delta_s\bar{z} - \Delta_n\bar{z} + \Delta_t\bar{z}. \tag{16}$$

On the right side, the terms are the change caused by selection, the change caused by the part of phenotype that is not associated with a transmitted predictor, and the change in the effects of the predictors during transmission.

Heritability and information

This section focuses on the amount of information that populations accumulate by selection, and the various processes that degrade or alter the transmission of that information. Some of the forms given here include the classic genetic measures of heritability as special cases. However, I do not emphasize those connections. Rather, I focus on general expressions given in terms of the full Price equation for total evolutionary change and based on predictors that may be chosen in any way. Different problems and goals will lead one to choose different sets of predictors or underlying causal schemes for phenotypes. The results here apply to any choice of predictors and causal scheme.

We start with eqn 15, the partition of phenotypic change into components

$$\Delta\bar{z} = \sum \Delta q_i z_i - \sum \Delta q_i \delta_i + \sum q'_i \Delta g_i.$$

The first term on the right side is the selection component, $\Delta_s\bar{z}$. From eqn 11, $\Delta_s\bar{z} = \alpha_z J$, where α_z changes scale between phenotype, z, and the gain in information by selection, J. Thus,

$$\Delta\bar{z} = \alpha_z J - \sum \Delta q_i \delta_i + \sum q'_i \Delta g_i.$$

Here, selection happens in the initial (parental) population, causing a gain in information, J. On the phenotypic scale, that gain in information is $\alpha_z J$. The remaining terms include processes that cause loss of information during transmission or cause other changes to phenotype.

The part of phenotype not transmitted

Start by assuming that the predictors and their effects do not change during transmission, $\Delta g_i = 0$. That assumption reduces total change to

$$\Delta_h \bar{z} = \alpha_z J - \sum \Delta q_i \delta_i$$

where $\Delta_h = \Delta_s - \Delta_n$ denotes the heritable component of selection, which is the total selection, Δ_s, minus the part of selective change that is not associated with predictors, Δ_n. The part not associated with predictors is not explicitly transmitted within the given model of phenotype.

The second term, $\sum \Delta q_i \delta_i$, has the general form (eqn 11) of the change in information

$$\sum \Delta q_i z_i = \alpha_z J,$$

which holds for any choice of z. Thus, letting $z \equiv \delta$, we obtain $\sum \Delta q_i \delta_i = \alpha_\delta J$. Putting this into the original expression yields

$$\Delta_h \bar{z} = \alpha_z J - \alpha_\delta J = (\alpha_z - \alpha_\delta) J.$$

The scale change terms, α, have the important additivity property that, in general, $\alpha_a + \alpha_b = \alpha_{a+b}$. Thus,

$$\alpha_z - \alpha_\delta = \alpha_{z-\delta} = \alpha_g,$$

because $g = z - \delta$. The expression for the change in phenotype, ignoring the change during transmission in the predictors and their effects, is

$$\Delta_h \bar{z} = \alpha_g J. \tag{17}$$

This expression is the information gain by selection, J, scaled by α_g, which relates the predicted phenotype, g, to the information accumulated by selection. Because $g = z - \delta$, we see that the amount of information transmitted is degraded by δ, the fraction of the phenotype, z, that is not explained by the predictors.

Change in transmitted components of phenotype

When we add back the remaining term to eqn 17, we obtain the full expression for phenotypic change as

$$\Delta \bar{z} = \alpha_g J + \sum q_i' \Delta g_i.$$

The last term is the change in the transmitted components of phenotype. From eqn 14, those components include changes in the predictors and changes in the effects of the predictors. A predictor's effect is its associated multiple regression coefficient. Multiple regression coefficients often change with context. On the one hand, the true underlying causal effect may change. On the other hand, our model of causality may not be exactly right, in which case shifting context will cause the assigned role of different predictors to change, even though the underlying causal effects of those predictors may not have changed.

Various approaches may be taken to evaluate the accuracy of the causal model, such as the stability of the predictor effects under changing context (Li, 1975). Typically, a better causal model has predictors with greater stability, shifting the components of total change more strongly to the $\alpha_g J$ information term. That increase in the information term is usually advantageous with respect to interpretation, because it is often hard to evaluate the meaning of changes in predictors and their effects in the second term.

Suppose, for example, that a significant component of phenotype is not explained by a stable set of predictors. Is the information accumulated by selection in the initial population lost during transmission because it is not associated with any transmissible component? Or, is that information transmitted by other predictors that are not included in our model? If the information does transmit by predictors not in our model, that information contributes to the second term with changing values of the predictors and their effects. Such changes are hard to interpret, because many different processes can potentially alter the predictors and their effects.

These fundamental equations of selection and evolution are, in a way, rather arbitrary, because they depend so strongly on the particular set of predictors that one chooses. What can we conclude? First, the equations are always true, and so give us a clear sense of the essential nature of selection, information and evolution. Second, a key part of understanding any problem concerns choosing the right set of predictors. Third, simple genetic models provide a good starting point in many cases, but rarely define a complete set of predictors and an accurate expression of causality. If one is able to model the causal scheme well, the analysis will often be simple and natural. I have emphasized a path analysis interpretation for the regression expressions, because path analysis emphasizes the choice of a good causal model.

Fisher's fundamental theorem

If we hold the predictors and their effects constant, then using eqn 17, the change in mean log fitness is

$$\Delta_h \bar{m} = \alpha_g J$$

for $m = g + \delta$. This expression for change in fitness, holding constant the predictors and their average effects, provides a generalization of Fisher's fundamental theorem of natural selection. Fisher used the presence or absence of allelic types as predictors, and the associated value of predicted fitness, g, as the genic value of fitness. With those definitions, the expression here is equivalent to Fisher's theorem. To translate back to the particular notation that Fisher used, one would translate the definitions for α_g and J into Fisher's forms. Frank (1997b) provides the tools for the translation, following Price (1972) and Ewens (1989). The point

here is that Fisher's theorem holds for any choice of predictors, as emphasized in Frank (1997b).

Causes of fitness

The expression $\Delta_s\bar{z} = \alpha_z J$ associates the accumulation of information by selection, J, with the selective component of phenotypic change. But that expression does not tell us why the association occurs. The phenotype may directly influence fitness. Alternatively, the phenotype may have no direct effect on fitness, but instead may be associated with some other process that influences fitness. A significant part of evolutionary analysis concerns evaluating the causes of fitness (Box 7).

We may analyse the causes of fitness in the same way that we analysed the causes of phenotype. We write our model, or hypothesis, for the causes of fitness as the regression equation

$$w = \phi + \pi z + \sum a_k y_k + \epsilon. \tag{18}$$

Here, ϕ is the baseline fitness when all other terms are zero; π is the average direct effect of the phenotype z on fitness, holding constant the other predictors of fitness; and a_k is the average effect of the other predictors of fitness, y_k. We may use any number of other predictors, and those predictors may be defined in any way, including factors in the model for phenotype. For example, predictors y_k can be alleles, nonlinear interactions between combinations of alleles, symbionts, maternal effects, cultural or environmental attributes, other phenotypes, phenotypes of neighbours and so on. The residual, ϵ, is the difference between the predicted value of fitness for a given set of predictors and the actual fitness.

A simple example

To study the role of different predictors of fitness, it is useful to reduce the model to just the direct effect, z, and one indirect effect, y, yielding

$$w = \phi + \pi z + ay + \epsilon.$$

In this partial regression equation, it is helpful to write out the regression coefficients in full notation to emphasize their interpretation. The partial regression coefficient $\pi = \beta_{wz\cdot y}$ is the average effect of z on w holding y constant, and $a = \beta_{wy\cdot z}$ is the average effect of y on w holding z constant, thus

$$w = \phi + \beta_{wz\cdot y} z + \beta_{wy\cdot z} y + \epsilon. \tag{19}$$

Condition for the increase of a phenotype by selection

Using the standard covariance form for selection based on eqn 6, the partial change in z caused by selection is

$$\bar{w}\Delta_s\bar{z} = \mathrm{Cov}(w, z),$$

which simply states that z increases by selection when it is positively associated with fitness. However, we now have the complication shown in eqn 19 that fitness also depends on another predictor, y. If we expand the covariance using the full expression for fitness in eqn 19, we obtain

$$\bar{w}\Delta_s\bar{z} = \beta_{wz\cdot y} V_z + \beta_{wy\cdot z}\mathrm{Cov}(y, z).$$

If we replace the covariance term by the product of a regression coefficient and a variance, $\beta_{yz}V_z$, we have

$$\Delta_s\bar{z} = (\beta_{wz\cdot y} + \beta_{wy\cdot z}\beta_{yz})V_z/\bar{w}. \tag{20}$$

The condition for the increase of z by selection is $\Delta_s\bar{z} > 0$. The same condition using the terms on the right side is

$$\beta_{yz}\beta_{wy\cdot z} + \beta_{wz\cdot y} > 0. \tag{21}$$

Let us use an abbreviated notation for the three terms

$$\begin{aligned}\beta_{yz} &= r\\ \beta_{wy\cdot z} &= B\\ \beta_{wz\cdot y} &= -C.\end{aligned}$$

The first term, $\beta_{yz} = r$, describes the association between the phenotype, z, and the other predictor, y. An increase in z by the amount Δz corresponds to an average increase of y by the amount (see Box 4)

$$\Delta y = r\Delta z.$$

The second term, $\beta_{wy\cdot z} = B$, describes the direct effect of the other predictor, y, on fitness, holding constant the focal phenotype, z. The third term, $\beta_{wz\cdot y} = -C$, describes the direct effect of the phenotype, z, on fitness, w, holding constant the effect of the other predictor, y.

Using the abbreviated notation, the condition for the increase in z by selection is

$$rB - C > 0.$$

The following sections interpret this condition in terms of three different biological scenarios.

Interactions between two species

I trace the effects of phenotype z in species A and phenotype y in species B on the fitness of types from species A (Frank, 1994, 1995c, 1997a). One may think of species B as an ecological partner that can influence the fitness of types from species A. Here, fitness always refers to effects on species A.

Unknown cause of association

I follow the path diagram in Fig. 1a. Increases in the phenotype, z, by an amount Δz, reduce fitness by $-C\Delta z$. Increases in the phenotype y directly benefit fitness by $B\Delta y$. The z and y phenotypes are associated by r,

although no specific cause is known. It may be that similar phenotypes tend to settle in the same area, or that a common environment of temperature and nutrients causes a phenotypic association. In any case, as z increases, the associated value of y changes on average by $\Delta y = r\Delta z$ and, equivalently, $B\Delta y = rB\Delta z$.

Tracing the pathways in Fig. 1a, an increase in the direct phenotype by Δz causes a change in fitness proportional to $(rB - C)\Delta z$, which is greater than zero when $rB - C > 0$. Thus, selection may favour an increase in z even though z directly decreases fitness, because the benefit from species B's phenotype, y, in proportion to rB, may outweigh the direct cost, $-C$.

Direct cause of association

Alternatively, suppose that the phenotype z directly enhances the vigour of its partners from species B. That direct effect of z on species B causes an increase in the benefit, y, that species B provides back to those with phenotype z. Fig. 1b shows this direct cause of y by z. The condition for z to be positively associated with fitness and to increase by selection remains $rB - C > 0$. However, the interpretation differs. In this case, z directly influences its neighbours' phenotype, y, rather than being associated with y by some unknown cause.

Body temperature

Suppose z is body temperature, which imposes a direct effect $-Cz$ on fitness. That direct cost may arise because body temperature raises the rate at which energy is used. Let y be speed of response to a challenge, such as a predator attack. Faster response provides a direct benefit, By. An unknown cause may associate temperature, z, and response rate, y, by an amount r (Fig. 1a). For example, sunshine may directly raise temperature and simultaneously increase response to attack by providing better visual opportunities. Alternatively, temperature, z, may directly raise response rate, y, by increasing the responsiveness of muscles (Fig. 1b). In either case, selection favours an increase in body temperature if $rB - C > 0$.

Social evolution and group selection

The phenotype z may be a costly altruistic behaviour that helps neighbouring individuals (Hamilton, 1970; Queller, 1992a, b; Frank, 1998). The direct effect on fitness is $-Cz$. Neighbours have phenotype y that provides a benefit, By, back to the original individual. An association, r, between z and y may arise in a variety of ways.

Some unknown cause may associate z and y (Fig. 1a). For example, shared cultural, environmental or genetic variation may cause related behaviour. Or a shared symbiont may cause an association. In general, any association in the predictors of phenotype will cause an association of phenotypic values.

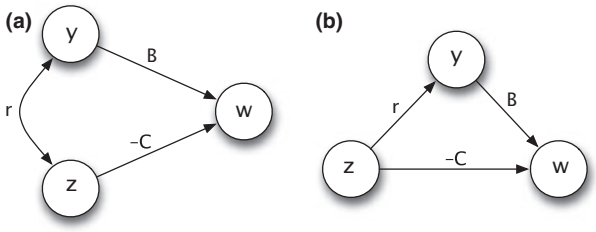

Fig. 1 Path diagrams for the effects of phenotype, z, and secondary predictor, y, on fitness, w. (a) An unknown cause associates y and z. The arrow connecting those factors points both ways, indicating no particular directionality in the hypothesized causal scheme. (b) The phenotype, z, directly affects the other predictor, y, which in turn affects fitness. The arrow pointing from z to y indicates the hypothesized direction of causality.

In other cases, the altruistic phenotype, z, may directly enhance neighbours' beneficial behaviour, y, in proportion to r (Fig. 1b). For example, the level of y in the neighbours may depend on the probability of the neighbours' survival. If an increase in z raises neighbours' survival in proportion to r, that increase in survival enhances the expression of the neighbours' behaviour, y, which has a beneficial effect on fitness of By.

Whether r arises from unknown causes (Fig. 1a) or from the direct effect of z on y (Fig. 1b), we can trace the effect of an increase in z on fitness. The condition for an increase in z to raise fitness is $rB - C > 0$.

In some cases, we may interpret the condition $rB - C > 0$ in terms of group selection (Hamilton, 1975). For example, z may measure individual restraint in the harvesting of nonrenewable resources (Frank, 1995b). Greater restraint reduces the direct benefit to the individual, because it means less resource harvested, with an effect on fitness of $-Cz$. Neighbours' phenotype, y, may be the average restraint among individuals in a local group with regard to harvesting nonrenewable resources.

Greater group restraint provides a benefit to all members of the group, including our focal individual, by providing greater local productivity through maintenance of nonrenewable resources. The benefit of group restraint on individual fitness is By. The association between an individual's phenotype, z, and the group phenotype, y, is r. Thus, when $rB - C > 0$, individual restraint evolves and provides a joint benefit to all group members. Here, the two predictors of fitness are individual behaviour, z, and average group behaviour, y. This type of group selection is just a special case of partitioning the causes of fitness, in which one of the predictors is a group attribute (Box 7).

Causal structure

All of these examples share a common causal structure. We are interested in the change in a phenotype, z,

caused by selection. Fitness depends on two predictors: the phenotype of interest, z, and another predictor, y. In all cases, the condition for the increase in z by selection is $rB - C > 0$. This condition is just the partition of the causes of fitness into two components. The direct effect on fitness of z is $-C$, and the direct effect of y is B. We multiply y by r to change the scale of the effect from y to z, because the net effect must be the relation between z and fitness, w.

We can see the logical relations and the units for the various scales by writing out the full notation

$$rB - C = \beta_{yz}\beta_{wy\cdot z} + \beta_{wz\cdot y}. \tag{22}$$

Box 3 shows that a regression coefficient, β_{xy}, has units $\Delta x/\Delta y$. Taking the terms of the above equation in order from left to right, the units are

$$\beta_{yz}\beta_{wy\cdot z} + \beta_{wz\cdot y} \equiv \frac{\Delta y}{\Delta z}\frac{\Delta w}{\Delta y} + \frac{\Delta w}{\Delta z} \equiv \frac{\Delta w}{\Delta z}. \tag{23}$$

The ratio $\Delta w/\Delta z$ is the change in fitness, w, per unit change in the phenotype, z. That ratio is the slope of fitness on phenotype. When the slope is positive, selection favours the increase of the phenotype. In any analysis of this sort, the term

$$r = \beta_{yz} = \frac{\Delta y}{\Delta z} \tag{24}$$

rescales changes of the secondary predictor, Δy, with respect to changes in the primary scale, Δz.

The key point is that $rB - C > 0$ simply partitions fitness into the direct effect of a phenotype plus the indirect effect through a secondary predictor. The true causal structure will, of course, frequently depend on multiple secondary causes, as in eqn 18. Multiple causes lead to an expanded expression for the increase of z caused by selection, $\Delta_s\bar{z}$, as

$$\sum r_i B_i - C > 0,$$

in which each r_i is the regression of y_i on z, and each B_i is the partial regression of w on y_i holding constant the other factors. One may also need to consider cascading causes or hidden factors in the sense of path analysis (Li, 1975). The simple expression $rB - C > 0$ should be thought of as a convenient example to illustrate the logic of partitioning the causes of fitness, or as the expression of simplified models that isolate two opposing processes.

In this section, I have analysed the partitioning of fitness. I have not discussed the partition of phenotype into components, $z = g + \delta$, where g is the sum of the predictors of phenotype. The amount of information accumulated by selection that can be transmitted depends on the slope of fitness, w, relative to the transmissible predictors of phenotype, g. If we think of g in terms of the genetic predictors of phenotype, then r can be interpreted as a genetic relatedness coefficient, and $rB - C > 0$ calls to mind Hamilton's rule from the theory of kin selection (Hamilton, 1970). The next article takes up the relations between kin selection and the general analysis of the causes of fitness and the causes of phenotype (Frank, 2013). A full evolutionary analysis also requires attention to other causes of change, $\Delta_t\bar{z}$, in eqn 16 (Frank, 1997b, 1998).

It is important to relate the causes of fitness to information, which is the ultimate scale for selection. Box 8 connects the partitions of fitness in this section to the expressions of information given earlier in this article.

Discussion

I first partitioned phenotype with respect to a set of hypothesized causes. I then partitioned fitness with respect to a different set of hypothesized causes. Finally, I placed those partitions of phenotype and fitness into a general expression for selection and evolutionary change. Those steps allowed me to express heritability, selection and evolutionary change in terms of causal components.

I also translated the standard expressions of selection and evolution, given in terms of regressions, covariances and variances, into expressions for the change in information. In my view, selection is best interpreted as the accumulation of information by populations (Frank, 2012a). Other evolutionary processes often cause a decay in the transmission of information. The information expressions allow one to read the equations of selection and evolution as if they were sentences. Those sentences express the fundamental relations between the causes of phenotypes and fitness and the consequences for the change in information by evolutionary processes.

I showed that the commonly used regressions coefficients in models of selection and evolution can be understood as coefficients for the change in scale with respect to the ultimate scale of information (Box 4). For example, the change in a phenotype caused by selection can be understood as a rescaling of the change in information accumulated by selection. Certain measures of heritability, often expressed as regression coefficients, are the change in the scaling of information from one phenotype to another. For example, a parent–offspring regression may describe the change in scale between parent and offspring phenotype with respect to the underlying information content in those phenotypes.

My extended development in terms of causal components and information may, at first, seem like a lot of technical complication. We are, after all, simply modelling selection, heritability and other widely studied evolutionary processes. Many models of those processes seem more direct and concise. My goal is to go beyond common calculations or common applica-

Box 8: Information and the causes of fitness

Changes caused by selection can always be related to the change in information accumulated by the population. For example, the change in phenotype caused by selection from eqn 11 is

$$\Delta_s \bar{z} = \alpha_z J,$$

where J is the change in information by selection, and α_z relates the scale of information to the scale of phenotype. We can examine the units of the scaling term

$$\alpha_z = \frac{\beta_{zw}}{\beta_{mw}},$$

which is the ratio of two regression coefficients (eqn 10). A regression coefficient, β_{zy}, has units $\Delta z/\Delta y$, when used as a scaling relation for changes in average values (Box 3). Thus, the units for the scaling relation, α_z, are

$$\frac{\beta_{zw}}{\beta_{mw}} \equiv \frac{\Delta z}{\Delta w}\frac{\Delta w}{\Delta m} = \frac{\Delta z}{\Delta m}.$$

The term Δm has units of change in log fitness. Changes in log fitness are equivalent to changes in information, J (eqn 8). To emphasize that J is a change in information, write the units on J as ΔI. Thus, the scaling factor

$$\alpha_z \equiv \frac{\Delta z}{\Delta I}$$

is the change in phenotype relative to the change in information.

One must learn to read the regression coefficients as scaling factors that change units. Once one learns to recognize the scale changes, and the key units such as information and phenotype, the fundamental equations can be read like a sentence. When analysing selection, I prefer information as the ultimate scale, because selection is the process by which populations accumulate information.

With that background, I present a long sentence to translate the causes of fitness into an expression for the change in information. Start with eqn 20 and divide both sizes by α_z, yielding

$$\Delta_s \bar{m} = \frac{\Delta_s \bar{z}}{\alpha_z} = \beta_{mw}\left(\frac{\beta_{wz\cdot y} + \beta_{wy\cdot z}\beta_{yz}}{\beta_{zw}}\right)\frac{V_z}{\bar{w}}.$$

The units are

$$\frac{\Delta_s \bar{z}}{\alpha_z} \equiv \Delta z\left(\frac{\Delta I}{\Delta z}\right) = \Delta I,$$

the change in information by selection. All of the regression coefficients in the prior equation change scales for the various terms, and we also have $V_z/\bar{w}$, which has units $\Delta z^2/\Delta w$. The net units of the long right side are ΔI, the change in information. The right side appears complex. But each term has a simple, readable meaning with respect to the effect of a predictor on fitness, and the scale changes required to transform those effects into the common units of information. To understand selection, we often need to decompose fitness and phenotypes into their component causes. Such decomposition requires that we combine all the components properly to recover the correct scale of analysis.

tions. The more abstract and exact models here provide a conceptual guide for understanding how selection actually works, how populations accumulate information, and how that information is transmitted or lost.

I have also traded the certainty of the standard models of genetics for the uncertainty that arises when we freely choose our predictors as causal hypotheses. In my view, the apparent certainty of genetics is often misleading. We know that many factors influence phenotypes in addition to the narrowly defined allelic types of genes. Traditionally, a specific extended model deals with each additional factor: cytoplasmic inheritance, nonlinear genetic interactions, maternal effects, social interactions and so on. By describing each of those aspects as a special situation, one ends up with a catalogue of special models.

The models here show how to think in general about a variety of causal structures. Those models are only as good as the particular hypothesized system of causality that we choose. But that is also true for genetic models and for every other model, whether or not we admit it openly. Here, I have traded the false sense that there are a few standard models for the more realistic view that one has to bring a good hypothesis to an analysis to get a good understanding of phenotypes and selection.

Hamilton (1970) made clear the central role of causal analysis in kin selection theory

> Considerations of genetical kinship can give a statistical reassociation of the [fitness] effects with the individuals that cause them.

The seemingly endless debates about kin selection arise from failure to recognize that the theory is ultimately a way of framing causal hypotheses (Frank, 1997b, 1998). The following article develops kin selection as a method of causal modelling.

Acknowledgments

National Science Foundation grant EF-0822399 supports my research.

References

Boyd, R. & Richerson, P.J. 1985. *Culture and the Evolutionary Process*. University of Chicago Press, Chicago.

Cavalli-Sforza, L.L. & Feldman, M.W. 1981. *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton University Press, Princeton, New Jersey.

Crespi, B.J. 1990. Measuring the effect of natural selection on phenotypic interaction systems. *Am. Nat.* **135**: 32–47.

Crespi, B.J. & Bookstein, F.L. 1989. A path-analytic model for the measurement of selection on morphology. *Evolution* **43**: 18–28.

Crow, J.F. & Kimura, M. 1970. *An Introduction to Population Genetics Theory*. Burgess, Minneapolis, Minnesota.

Dawkins, R. 1976. *The Selfish Gene*. Oxford University Press, New York.

Ewens, W.J. 1989. An interpretation and proof of the fundamental theorem of natural selection. *Theor. Popul. Biol.* **36**: 167–180.

Falconer, D.S. & Mackay, T.F.C. 1996. *Introduction to Quantitative Genetics*, 4th edn . Longman, Essex, England.

Fisher, R.A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edinb.* **52**: 399–433.

Fisher, R.A. 1930. *The Genetical Theory of Natural Selection*. Clarendon, Oxford.

Frank, S.A. 1986. Hierarchical selection theory and sex ratios I. General solutions for structured populations. *Theor. Popul. Biol.* **29**: 312–342.

Frank, S.A. 1994. Genetics of mutualism: the evolution of altruism between species. *J. Theor. Biol.* **170**: 393–400.

Frank, S.A. 1995a. George Price's contributions to evolutionary genetics. *J. Theor. Biol.* **175**: 373–388.

Frank, S.A. 1995b. Mutual policing and repression of competition in the evolution of cooperative groups. *Nature* **377**: 520–522.

Frank, S.A. 1995c. The origin of synergistic symbiosis. *J. Theor. Biol.* **176**: 403–410.

Frank, S.A. 1997a. Models of symbiosis. *Am. Nat.* **150**: S80–S99.

Frank, S.A. 1997b. The Price equation, Fisher's fundamental theorem, kin selection, and causal analysis. *Evolution* **51**: 1712–1729.

Frank, S.A. 1998. *Foundations of Social Evolution*. Princeton University Press, Princeton, New Jersey.

Frank, S.A. 2012a. Natural selection. IV. The Price equation. *J. Evol. Biol.* **25**: 1002–1019.

Frank, S.A. 2012b. Natural selection. V. How to read the fundamental equations of evolutionary change in terms of information theory. *J. Evol. Biol.* **25**: 2377–2396.

Frank, S.A. 2013. Natural selection. VII. history and interpretation of kin selection theory. *J. Evol. Biol.* **26**: (in press).

Hamilton, W.D. 1970. Selfish and spiteful behaviour in an evolutionary model. *Nature* **228**: 1218–1220.

Hamilton, W.D. 1975. Innate social aptitudes of man: an approach from evolutionary genetics. In: *Biosocial Anthropology* (R. Fox, ed), pp. 133–155. Wiley, New York.

Heisler, I.L. & Damuth, J. 1987. A method for analyzing selection in hierarchically structured populations. *Am. Nat.* **130**: 582–602.

Kingsolver, J.G. & Schemske, D.W. 1991. Path analyses of selection. *Trends Ecol. Evol.* **6**: 276–280.

Lande, R. & Arnold, S.J. 1983. The measurement of selection on correlated characters. *Evolution* **37**: 1212–1226.

Li, C.C. 1975. *Path Analysis*. Boxwood, Pacific Grove, California.

Lynch, M. & Walsh, B. 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland, Massachusetts.

Price, G.R. 1972. Fisher's 'fundamental theorem' made clear. *Ann. Hum. Genet.* **36**: 129–140.

Price, G.R. 1995. The nature of selection. *J. Theor. Biol.* **175**: 389–396.

Queller, D.C. 1992a. A general model for kin selection. *Evolution* **46**: 376–380.

Queller, D.C. 1992b. Quantitative genetics, inclusive fitness, and group selection. *Am. Nat.* **139**: 540–558.

Scheiner, S.M., Mitchell, R.J. & Callahan, H.S. 2000. Using path analysis to measure natural selection. *J. Evol. Biol.* **13**: 423–433.

Taylor, P.D. & Frank, S.A. 1996. How to make a kin selection model. *J. Theor. Biol.* **180**: 27–37.

Received 29 August 2012; accepted 4 November 2012

JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.12131

REVIEW

Natural selection. VII. History and interpretation of kin selection theory*

S. A. FRANK

Department of Ecology and Evolutionary Biology, University of California, Irvine, CA, USA

Keywords:

altruism;
causal analysis;
cooperation;
group selection;
inclusive fitness;
population genetics;
quantitative genetics;
social evolution.

Abstract

Kin selection theory is a kind of causal analysis. The initial form of kin selection ascribed cause to costs, benefits and genetic relatedness. The theory then slowly developed a deeper and more sophisticated approach to partitioning the causes of social evolution. Controversy followed because causal analysis inevitably attracts opposing views. It is always possible to separate total effects into different component causes. Alternative causal schemes emphasize different aspects of a problem, reflecting the distinct goals, interests and biases of different perspectives. For example, group selection is a particular causal scheme with certain advantages and significant limitations. Ultimately, to use kin selection theory to analyse natural patterns and to understand the history of debates over different approaches, one must follow the underlying history of causal analysis. This article describes the history of kin selection theory, with emphasis on how the causal perspective improved through the study of key patterns of natural history, such as dispersal and sex ratio, and through a unified approach to demographic and social processes. Independent historical developments in the multivariate analysis of quantitative traits merged with the causal analysis of social evolution by kin selection.

As is often the case, once a topic has become in vogue, its name ceases to have meaning …

(Lazar & Birnbaum, 2012).

Introduction

Ideas are embedded in their history and language. Hamilton's (1970) theories of inclusive fitness and kin selection are good examples. As understanding deepened, the original ideas transformed into broader concepts of selection and evolutionary process. With that generalization, the initial language that remains associated with the topic has become distorted. The confused language and haphazard use of incorrect historical context have led to significant misunderstanding and meaningless argument.

Correspondence: Steven A. Frank, Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697-2525, USA. Tel.: +1 949 824 2244; fax: +1 949 824 2181; e-mail: safrank@uci.edu

*Part of the Topics in Natural Selection series. See Box 1.

Current understanding transcends the initial interpretation of 'kin selection,' which attaches to some notion of similarity by descent from a recent common ancestor. The other candidate phrases, such as 'inclusive fitness' or 'group selection,' also have problems. We are left with a topic that derives from those antecedent notions and clearly has useful application to those biological puzzles. At the same time, the modern understanding of altruism connects to the analysis of selection on multiple characters, to interactions between different species and to the broadest generalizations of the theory of natural selection.

A good scholarly history of kin selection and its descendants has yet to be written. Here, I give a rough historical outline in the form of a nonmathematical narrative (see Box 2). I describe the history from my personal perspective. Because I worked actively on the subject over several decades, I perceive the history by the ways in which my own understanding changed over time. Box 3 highlights other perspectives and key citations.

Box 1: Topics in the theory of natural selection

This article is part of a series on natural selection. Although the theory of natural selection is simple, it remains endlessly contentious and difficult to apply. My goal is to make more accessible the concepts that are so important, yet either mostly unknown or widely misunderstood. I write in a nontechnical style, showing the key equations and results rather than providing full derivations or discussions of mathematical problems. Boxes list technical issues and brief summaries of the literature.

Early history

According to Darwin (1859), natural selection favours traits that enhance individual reproduction. Puzzles arise when traits reduce individual reproduction while providing aid to others. In this context, Fisher (1930) discussed the problem of warning colouration in mimicry:

> [D]istastefulness … is obviously capable of giving protection to the species as a whole, through its effect upon the instinctive or acquired responses of predators, yet since any individual tasted would seem almost bound to perish, it is difficult to perceive how individual increments of the distasteful quality, beyond the average level of the species, could confer any individual advantage.

Fisher then gave one possible solution:

> [W]ith gregarious larvae, the effect will certainly be to give the increased protection especially to one particular group of larvae, probably brothers and sisters of the individual attacked. The selective potency of the avoidance of brothers will of course be only half as great as if the individual itself were protected; against this is to be set the fact that it applies to the whole of a possibly numerous brood. … The ideal of heroism has been developed among such peoples considerably beyond the optimum of personal advantage, and its evolution is only to be explained, in terms of known causes, by the advantage which it confers, by repute and prestige, upon the kindred of the hero.

Haldane (1955) restated Fisher's argument. Williams & Williams (1957) presented a specific and limited analysis, which they applied to altruism in social insects. Some rather vague models considered more diffuse forms of genetic similarity created by population structuring of groups (Haldane, 1932; Wright, 1945). These analyses hinted at a more general concept. However, none of them expressed the deeper principle in a clear and convincing way.

Hamilton's rule

Hamilton (1963, 1964a,b) initiated modern approaches by expressing a rule for the increase in an altruistic behaviour

$$rB - C > 0. \quad (1)$$

Here, C is the cost to an actor for performing the altruistic behaviour, B is the benefit gained by a recipient of the altruistic act and r measures the relatedness between actor and recipient. The idea is that the actor pays a personal cost in reduced reproduction for helping another individual, the recipient gains increased reproduction from the altruistic act and the relatedness translates the recipient's enhanced reproduction back to the actor's valuation of that extra reproduction. When the benefit back to the actor, rB, is greater than the actor's direct cost, C, then the behaviour is favoured by natural selection.

Hamilton figured out how to measure r by analysing population genetic models. With the methods Hamilton

Box 2: Scope

> Quoting from Fawcett & Higginson (2012): Most research in biology is empirical, yet empirical studies rely fundamentally on theoretical work for generating testable predictions and interpreting observations. Despite this interdependence, many empirical studies build largely on other empirical studies with little direct reference to relevant theory, suggesting a failure of communication that may hinder scientific progress. … The density of equations in an article has a significant negative impact on citation rates, with papers receiving 28% fewer citations overall for each additional equation per page in the main text. Long, equation-dense papers tend to be more frequently cited by other theoretical papers, but this increase is outweighed by a sharp drop in citations from non-theoretical papers.

This article contains no equations beyond a few summary expressions. I do not write for other theoreticians. I do not attempt to be comprehensive. Rather, I try to evoke some lines of thought that I believe will be helpful to scientists who want to know about the theory.

A rough idea about the theory aids empirical study. It also helps to cope with the onslaught of theoretical articles. Those theoretical articles often claim to shift the proper framing of fundamental issues. The literature never seems to come to a consensus.

Here, I attempt to translate a few of the key points into nonmathematical summaries. Such translation necessarily loses essential components of understanding. Yet it seems worthwhile to express the main issues in way that can be understood by a wider audience. Refer to my earlier work for mathematical aspects of the theory and for citations to the technical literature (Frank, 1997a,b, 1998, 2012a,b,c, 2013).

Box 3: Literature

For each topic related to kin selection theory, I list a small sample of key articles and reviews. This limited space does not allow comprehensive coverage or commentary on the particular articles, but should provide an entry into the extensive literature and the range of opinions.

Several reviews follow Hamilton's perspective (Alexander, 1974; Dawkins, 1979; Michod, 1982; Grafen, 1985; Lehmann & Keller, 2006; Wenseleers, 2006; Dugatkin, 2007; Bourke, 2011; Gardner *et al.*, 2011). An associated literature emphasizes the problem of sociality and sterile castes in insects, with additional commentary on general aspects of the theory (Wilson, 1971; West-Eberhard, 1975; Trivers & Hare, 1976; Andersson, 1984; Brockmann, 1984; Alexander *et al.*, 1991; Bourke & Franks, 1995; Queller & Strassmann, 1998; Foster *et al.*, 2006).

Kin selection theory has been applied to a wide range of biological problems. Here, I can list only a few general overviews. Those overviews give a sense of the scope but do not include many significant applications (Trivers, 1985; Maynard Smith & Szathmáry, 1995; Crespi, 2001; Michod & Roze, 2001; West *et al.*, 2007; Burt & Trivers, 2008; West, 2009; Davies *et al.*, 2012).

The strongest criticisms arose from population genetics. The main issues concern how the specifics of genetics can vary from case to case and alter the outcome of selection, and how the full analysis of dynamics may provide an essential, deeper perspective on evolutionary process (Uyenoyama *et al.*, 1981; Uyenoyama & Feldman, 1982; Karlin & Matessi, 1983; Kerr *et al.*, 2004; Nowak *et al.*, 2010).

Kin selection theory has a long association with debates about units and levels of selection. I give a very short listing, because that topic is beyond my scope (Lewontin, 1970; Dawkins, 1982; Keller, 1999; Okasha, 2006). The related topic concerning group selection does fall within my scope (Wade, 1978; Uyenoyama & Feldman, 1980; Wilson, 1983; Grafen, 1984; Nunney, 1985; Wade, 1985; Heisler & Damuth, 1987; Queller, 1992a; Dugatkin & Reeve, 1994; Soltis *et al.*, 1995; Sober & Wilson, 1998; Henrich, 2004; Traulsen & Nowak, 2006; West *et al.*, 2008; Leigh, 2010).

The merging of kin selection theory with quantitative genetics and multivariate analyses of selection follows various lines of development (Cheverud, 1984; Queller, 1992b; Wolf *et al.*, 1998, 1999; Bijma & Wade, 2008; McGlothlin *et al.*, 2010; Wolf & Moore, 2010). Advanced aspects of the theory and new directions of theoretical development continue to appear (Rousset, 2004; Grafen, 2006; Taylor *et al.*, 2006; Gardner *et al.*, 2007; Fletcher & Doebeli, 2009).

used in those early papers, he was only able to give a rough description for the proper measure for relatedness. His initial expression in terms of the genetic correlation between actor and recipient was in the right direction and was later refined by Hamilton (1970).

Hamilton also defined a new and more general notion of fitness, which he called 'inclusive fitness'. Instead of counting the number of offspring by an individual, inclusive fitness makes a more extensive calculation of how phenotypes influence the transmission of genes from one generation to the next. In the inclusive fitness interpretation, eqn 1 describes the changes in direct and indirect reproduction associated with each change in behaviour. Thus, the inclusive fitness of a particular behavioural act is the indirect reproductive gain through the recipient, B, multiplied by the relatedness, r, minus the loss in direct reproduction, C. The relatedness, r, measures the genetic discount of substituting the reproduction of the recipient in place of the reproduction of the actor.

All of the direct and indirect fitness changes are assigned to the actor. This assignment of the different pathways of genetic transmission to the actor associates all fitness changes with the behaviour that caused those changes. Hamilton viewed this inclusive assignment of all genetic consequences to the original causal behaviour as the key to his approach. The proper way of doing the calculation was the major theoretical advance. Hamilton (1970) emphasized this causal perspective:

> Considerations of genetical kinship can give a statistical reassociation of the [fitness] effects with the individuals that cause them.

Hamilton's 1970 paper also greatly advanced the analysis by using Price's equation (Price, 1970). With this method, Hamilton established the correct measure for genetic similarity, r, between actor and recipient. That measure turned out to be the regression coefficient of the recipient's genotype in relation to the actor's genotype. That regression is the value needed to establish a proper measure of inclusive fitness. Hamilton noted that the regression was the exact measure only in his particular models. He argued that the regression measure would extend, at least approximately, to a wide variety of other assumptions.

Roughly speaking, r translates the actor's deviation in gene frequency from the population average to the recipient's deviation in gene frequency from the population average. Thus, recipient's reproduction has consequence for gene frequency change that is r multiplied by what the same fitness increment in the actor would cause to gene frequency change (Grafen, 1985; Frank, 1998, p. 49).

Comparison of different approaches

Already by 1975, different lines of thought on altruism had developed. Hamilton (1975, pp. 336–337) gave his opinion:

> The usefulness of the 'inclusive fitness' approach to social behaviour (i.e. an approach using criteria like [$rB - C > 0$]) is that it is more general than the 'group selection', 'kin selection' or 'reciprocal altruism' approaches and so provides an overview even where regression coefficients [r] and fitness effects [B and C] are not easy to estimate or specify.

For Hamilton, his rule was a way to clarify biological understanding and to develop qualitative hypotheses about the causes of adaptation. Hamilton did not think of his rule as replacing the way in which one did calculations for models of population genetics. Indeed, he always considered the classical population genetics theory as the primary truth. He then evaluated his own methods in light of how well they could capture, in a simple way, the complexities of the underlying genetic models. Following that historical line of thought, the subject subsequently split into two lineages.

On the biological side, Hamilton's perspective completely changed the way people approach a great variety of key problems, ranging from social insects to parasite virulence to bacterial competition to the evolutionary history of 'individuals' to the historical tendency for an increase in biological complexity. Almost everyone agrees that this was a revolutionary change in biological thought and that it derived from Hamilton's work.

At the same time, heated debates arose about theoretical interpretations and mathematical details. We see in Hamilton's (1975) quote the different competing phrases and associated theoretical perspectives. The ongoing debates have grown ever more fierce rather than settling out to a common perspective.

In essence, I think there is almost no disagreement about how evolutionary process shapes biological characters. No matter the perspective, when faced with the same biological problem, all of the different approaches usually arrive at roughly the same predictions about how evolution shapes characteristics. Yet, in spite of that agreement, the arguments persist about whether one should call the underlying process 'group selection' or 'kin selection' or 'inclusive fitness' or 'population genetics' or whatever else is being promoted.

Clearly, we need to understand more than just the predicted outcomes: we must also understand the underlying causal processes. So something is at stake here. But what exactly? The best way to understand that question is through the historical development of the subject. So, let us continue with Hamilton's (1975) article and subsequent work.

Almost everything that one would reasonably want to say about group selection in relation to inclusive fitness or kin selection is in the following quotes from Hamilton (1975). I quote in full because there has been much controversy and misunderstanding about these issues. It helps to read Hamilton's perspective, given long before the current participants in the debates fully developed their views on the subject.

> As against 'group selection' it [inclusive fitness] provides a useful conceptual tool where no grouping is apparent – for example, it can deal with an ungrouped viscous population where, owing to restricted migration, an individual's normal neighbours and interactants tend to be his genetical kindred.

In other words, inclusive fitness is more general. Group selection is just a case in which the positive association, r, arises from clearly defined aspects of groups. In cases for which groups are not easily delineated, the same underlying inclusive fitness approach still holds. Continuing

> Because of the way it was first explained, the approach using inclusive fitness has often been identified with 'kin selection' and presented strictly as an alternative to 'group selection' as a way of establishing altruistic social behaviour by natural selection (Maynard Smith, 1964; Lewontin, 1970). However the foregoing discussion shows that kinship should be considered just one way of getting positive regression [r] of genotype in the recipient, and that it is this positive regression that is vitally necessary for altruism. Thus, the inclusive-fitness concept is more general than 'kin selection'.

Hamilton preferred to reserve 'kin selection' for cases in which the positive regression, r, comes from interactions between individuals that we would commonly describe by terms of kinship, such as cousins. I will later argue against Hamilton's use of 'inclusive fitness' as the ultimate causal view. I will end up using 'kin selection' as the label for a wide variety of processes, because of the lack of a better alternative. However, in 1975, Hamilton's view made sense. For now, it is useful to read what Hamilton said to understand his perspective on the different framings for causal process. Continuing

> Haldane's [(1932)] suggestion about tribe-splitting can be seen in one light as a way of increasing intergroup variance and in another as a way of getting positive regression in the population as a whole by having the groups which happen to have most altruists divide most frequently. In this case, the altruists are helping true relatives. But in the assortative-settling model, it obviously makes no difference if altruists settle with altruists because they are related (perhaps never having parted from them) or because they recognize fellow altruists as such, or settle together because of some pleiotropic effect of the gene on habitat preference. If we insist that group selection is different from kin selection the term should be restricted to situations of assortation definitely not involving kin. But it seems on the whole preferable to retain a more flexible use

> of terms; to use group selection where groups are clearly in evidence and to qualify with mention of 'kin' …, 'relatedness' or 'low migration' (which is often the cause of relatedness in groups), or else 'assortation', as appropriate. The term 'kin selection' appeals most where pedigrees tend to be unbounded and interwoven, as is so often the case with humans.

The point is that the different labels serve only to help identify the cause of association, r. The underlying evolutionary process should be understood with respect to $rB - C > 0$, even when it is difficult in practice to calculate directly the different terms in Hamilton's inequality. Hamilton favoured analysing complex biological problems with population genetic models, then interpreting those models in terms of the simple causal framework captured by the r, B and C components of his rule.

Using Hamilton's theory to solve problems

Limitation of Hamilton's theory in practical applications

Hamilton (1972) used his rule to develop various qualitative hypotheses about social insect evolution. When Hamilton analyzed kin interactions in quantitative models of sex ratios (Hamilton, 1967) and dispersal (Hamilton & May, 1977), he first made his mathematical calculations with genetic or game theory models, then made *post hoc* interpretations of the quantitative results in terms of interactions between kin. Hamilton never used inclusive fitness theory or Hamilton's rule to solve for the quantitative phenotype favoured by selection. Hamilton (1975, p. 337) noted that

> Although correlation between interactants is necessary if altruism is to receive positive selection, it may well be that trying to find regression coefficients is not the best analytical approach to a particular model. Indeed, the problem of formulating them exactly for sexual models proves difficult (Chapter 2). One recent model that makes more frequent group extinction the penalty for selfishness (or lack of altruism) has achieved rigorous and striking conclusions without reference to regression or relatedness (Eshel, 1972). But reassuringly, the conclusions of both this and another similar model (Levins, 1970) are of the general kind that consideration of regression leads us to expect. The regression is due to relatedness in these cases, but classified by approach these were the first working models of group selection.

In 1979, I took Hamilton's graduate seminar at the University of Michigan. I inherited Hamilton's interest in fig wasp sex ratios and the idea that one could develop models of kin interactions and sex ratios using the Price equation. At that time, Hamilton was losing interest in working on such problems, and I was left with the last seeds of his insight on this subject. As I pursued my empirical studies of fig wasp sex ratios, I also tried to learn how one could develop more realistic models of sex ratios with complex kin interactions.

Using kin selection theory to analyse models of kin interactions

At first, I had Hamilton's doubts about using kin selection theory directly to solve problems. Instead, the method in those days was to solve the problem with population genetics, and then try to interpret the resulting predictions in terms of kin interactions. The biological interactions from my empirical studies led to horrendously complex population genetic analyses (Frank, 1983b). With great effort, I could solve some of the problems. I repeatedly found that the *post hoc* interpretations in terms of kin selection were simple and easy to understand. For example, the value to a mother of an extra son is devalued by the mother's genetic relatedness to the competitors of her sons (Frank, 1985b, 1986b,c).

This underlying simplicity in the exact results of population genetics led me to try Hamilton's suggestion about using the Price equation to model sex ratios. Eventually, I found a simple Price equation method to obtain the same results as the complex genetic approach, when analysing commonly used assumptions. The Price equation method also gave results that were much more general than the population genetic methods. In a Price equation analysis, it was easy to follow the causal processes during the derivation, and it was easy to interpret the final results in terms of the biology. The method was like reading sentences from a book in which the biological processes of competition, cooperation and kin interactions were written in the clearest and most direct manner.

The generality of my solutions, derived directly in terms of kin interactions, improved through a series of papers on sex ratios and dispersal (Frank, 1985a,b, 1986a,b,c, 1987). At first, it did not make sense to twist the results for those biological applications into a form that looked like Hamilton's rule, $rB - C > 0$. The results did not appear in terms of simple costs and benefits. Consequently, I gave little thought to Hamilton's rule, and instead thought only of how kin interactions shaped the evolution of interesting biological characters.

Solving problems: dispersal example

A model of dispersal illustrates how the application of kin selection theory changed through the 1970s and 1980s. In the 1970s, Hamilton (1978, 1979) became interested in wing polymorphisms among insects. For

example, some of the parasitic wasps that live in figs have two male morphs. A wingless male stays within its natal fig to compete for nearby mates. A winged male leaves to search for mates in other figs. Hamilton guessed that the dispersers must often die before finding another fig containing potential mates. With such an extreme cost of dispersal, why would an individual develop to disperse rather than stay and try to mate locally?

Hamilton & May (1977) realized that competition between kin may explain dispersal even when the cost is very high. A male that stays and outcompetes brothers replaces the brother's sperm with his genetically similar sperm, causing little gain in the net success of their shared genotype. By contrast, any success of a male disperser against nonrelated males provides full benefit to the fitness gain of his genotype.

Although Hamilton & May (1977) recognized the kin selection processes involved, they did not use kin selection theory as a method to analyse the problem. Instead, they formulated a simple ecological model and then solved for the evolutionarily stable strategy (ESS). After obtaining the result, they then gave a *post hoc* interpretation in terms of kin interactions.

In 1986, I solved this problem in a much simpler and more general way by using kin selection theory as a method of analysis (Frank, 1986a). This history of the kin selection approach sets the context for the modern understanding of the theory. The following is a slightly modified summary from Frank (1998, Section 7.2).

Hamilton & May's ESS model

Hamilton & May (1977) assumed that a habitat has a large number of discrete sites. In each year, the parents die after producing offspring. Each offspring has a trait that determines the probability, d, that it disperses from its natal patch. Those that stay at home, with probability $1 - d$, compete for one of N available breeding sites. Dispersers die with probability c, and with probability $1 - c$ they find a patch in which to compete for breeding. All sites are occupied in the simple model discussed here. Hamilton & May analysed the case in which one breeding site ($N = 1$) is available in each patch. In an asexual model, Hamilton & May found the ESS dispersal fraction, d, to be

$$d^* = \frac{1}{1+c}, \tag{2}$$

in which c, the cost of dispersal, varies from just above zero on up to nearly one. Interestingly, even if the cost of dispersal is high and the chance of surviving the dispersal phase is low, nearly half of the offspring still disperse.

In a second model, Hamilton & May analysed dispersal in a sexual species. Sexual reproduction raised analytical difficulties, because some inbreeding is likely if mating takes place within the local patch. Their method of analysis did not easily handle inbreeding, so they assumed that all males disperse before mating, and a fraction d of females disperse. With that set-up, they found that the ESS dispersal fraction for females is

$$d^* = \frac{1-2c}{1-2c^2}, \tag{3}$$

in which the dispersal rate is zero when $c \geq 1/2$. Hamilton & May understood the role of kin selection. The dispersal rate is lower in the sexual compared with the asexual model because the genetic relatedness of competitors within the natal patch was less in the sexual model. Less kin competition reduces the benefit of dispersing to avoid competition with kin, and so lowers the dispersal rate favoured by selection. Beyond that intuitive *post hoc* interpretation, kin selection theory played no role in the analysis.

Motro's population genetic analysis

Motro (1982a,b, 1983) analysed a fully dynamic population genetic model for the same problem. His complex analysis, spread over three articles, covered the same biological assumptions as Hamilton & May, with a few minor extensions. The length and complexity of Motro's work arose from the detailed population genetic analysis, as opposed Hamilton & May's relatively simple ESS methods.

Motro found that, in an asexual model, his result agreed with Hamilton & May's expression given in eqn 2. Motro obtained two additional results. First, in a sexual model in which the mother controls the dispersal trait of offspring, the same result arises as for the asexual model, as in eqn 2. By contrast, when Motro tried to match the assumptions of Hamilton & May for offspring control of phenotype, he obtained

$$d^* = \frac{1-4c}{1-4c^2}, \tag{4}$$

which differs from Hamilton & May's result in eqn 3. Motro drew two conclusions. First, in sexual models, the equilibrium depends on whether offspring phenotype is controlled by the mother or by the offspring. Second, under offspring control, explicit population genetic models failed to confirm Hamilton & May's result. Motro attributed the mismatch to a failure of the simplified ESS method when compared with his exact and rigorous population genetic techniques.

Analysis by kin selection

I analysed this dispersal problem by kin selection to understand the different results of Hamilton & May and Motro (Frank, 1986a). Box 4 shows the expression for fitness and some technical details.

Following Maynard Smith (1982), the standard ESS approach is to find a local maximum of fitness with

Box 4: Fitness expression for dispersal

In the dispersal model, fitness w depends on the dispersal phenotypes at the three different scales: a focal individual, d; the average dispersal probability in the focal individual's patch, d_p; and the population average, $\bar{d}$, yielding

$$w(d, d_p, \bar{d}) = (1 - d)\, p(d_p) + d(1 - c)\, p(\bar{d}). \qquad (11)$$

When an individual remains at home, with probability $1 - d$, its expected success, $p(d_p)$, depends on the average dispersal fraction in its patch, d_p. When an individual disperses, with probability d, its probability of landing in a new patch is $1 - c$, and its expected success in the new patch, $p(\bar{d})$, depends on the average dispersal probability in the population, $\bar{d}$. The expected success expressions, $p(d_p)$ and $p(\bar{d})$, can be written as

$$p(\alpha) = \frac{1}{1 - \alpha + \bar{d}(1 - c)},$$

in which one can use either $\alpha \equiv d_p$ or $\alpha \equiv \bar{d}$. The denominator is proportional to the number of competitors for breeding in a patch, and so the reciprocal is proportional to the expected success per individual in that patch.

When analysing the fitness maximum with respect to phenotype, one takes the derivative of w in eqn 11 with respect to d. That differentiation leads to terms in which one has the derivative of d_p with respect d. Such a term would require specifying how the average phenotype of neighbours in a natal patch, d_p, changes with respect to the phenotype of a focal individual, d. With kin interactions, that relationship could be complex, because the focal individual's phenotype, d, would be correlated with the average phenotype of its neighbours, d_p, through the genetic similarity among neighbours.

The lack of clarity about the slope of neighbour (or group) phenotype with respect to the focal individual's phenotype initially led to the abandonment of the simple maximization method for ESS analysis when genetic relatives interacted. In 1986, when I first analysed this dispersal problem, I published a Price equation method of analysis. I also began to see that the simple ESS maximization method worked when one simply replaced the derivative of neighbour phenotype on actor phenotype by the coefficient of relatedness. One could see the equivalence by matching up terms in a Price equation analysis with what one obtained by differentiating fitness with respect to phenotype and expanding out the terms. However, in the 1980s, I was not certain about how to justify using the simple maximization approach, so I published only Price equation analyses. Later, in 1996, with the help of Peter Taylor's deep understanding and elegant analysis, we published the general maximization method (Taylor & Frank, 1996).

respect to phenotype. When the amount of genetic variation for the phenotype is small, that local maximum is an ESS. At the time, it was generally thought that this ESS maximization method would not work with kin interactions, and so was not used to analyse problems of kin selection (Box 4).

I compared my Price equation analysis for the change in fitness with the terms obtained by an ESS maximization approach. The potentially difficult term in the ESS analysis is the slope (derivative) of average group phenotype with respect to individual phenotype. That slope from the ESS analysis always matched a regression coefficient in the Price equation. Under common assumptions, that regression coefficient is exactly the coefficient of relatedness of an individual to its neighbours. Thus, one can often replace the slope of group phenotype with respect to individual phenotype by the coefficient of relatedness from kin selection theory, r (Box 4).

All of that may sound a bit complicated, but in practice it is quite easy. Take the derivative of fitness with respect to the individual phenotype; set to r the slope of average group phenotype with respect to individual phenotype; and solve for a local maximum to obtain

$$d^* = \frac{r - c}{r - c^2}, \qquad (5)$$

which is the result reported in Frank (1986a) and discussed in more detail in Frank (1998, Section 7.2).

Kin selection simplifies and generalizes prior results

Motro and Hamilton & May assumed that only one female could breed in each patch. They made that assumption because their methods did not allow them to analyse the more complicated situation in which multiple females bred in each patch. To compare eqn 5 to the prior models, let us first follow that earlier assumption of one breeding female per patch.

If the organism is asexual, then in each patch the candidates for dispersal – the offspring of the single asexual mother – are related by a coefficient of one, $r = 1$, and eqn 5 reduces to Hamilton & May's result in eqn 2. If the organism is sexual, and offspring phenotype is controlled by the mother, then $r = 1$, and we again have eqn 2. The reason $r = 1$ with maternal control is that it is the relatedness of the individual that controls phenotype to the average phenotype of its patch that matters. With only one breeding female, the mother's phenotype and the average phenotype in the patch are the same, and $r = 1$.

If, by contrast, offspring control their own phenotype, then relatedness among competitors is $r = 1/2$, because competitors on a patch are outbred full siblings. With $r = 1/2$ in eqn 5, we recover the second result of Hamilton & May in eqn 3.

Hamilton & May assumed that the mother mated only once, so that siblings are related by $r = 1/2$. By contrast,

Motro implicitly assumed that the mother mated several times and that offspring in a patch were only half sibs, so in his model $r = 1/4$. Using that value of r in the general solution of eqn 5 yields Motro's result in eqn 4. Thus, the single kin selection model of eqn 5 explains the parent–offspring conflict and the difference between Motro's analysis and Hamilton & May's model.

With the kin selection model, we are not limited to one breeding site per patch, or to an outbreeding system. Rather, we can treat r as a parameter and express the ESS dispersal fraction in terms of the coefficient of relatedness. Higher relatedness increases dispersal. The reason is that a genotype competing with close relatives gains little by winning locally against its relatives. Even a small chance of successful migration and competition against nonrelatives can be favoured.

Discussion of the new kin selection methods of analysis

In retrospect, the kin selection analysis of dispersal seems simple and obvious. Yet, at the time, Hamilton had not been able to use kin selection theory to analyse the problems of sex ratio and dispersal that interested him. In the theoretical analysis of phenotypes, kin selection was only a *post hoc* method of interpreting results obtained by other means. The understanding of process was sufficiently confused that Motro could write three articles criticizing the Hamilton & May model, arguing that only a formal population genetic analysis could give the correct results. In fact, Motro simply made different assumptions about whether the dispersal phenotype was controlled by the parent or the offspring and about whether females mated once or many times. Those distinctions become obvious from the simpler and more general perspective of the kin selection analysis of phenotype.

The solution in eqn 5 did not have any obvious connection to Hamilton's rule. Similar kin selection analyses of sex ratio models also did not connect in any clear way to Hamilton's rule (Frank, 1985a,b, 1986a,b,c, 1987). At that time, I had concluded that Hamilton's rule was not useful for solving realistic problems. In application, nothing like Hamilton's rule appeared. It turned out that I was wrong about the generality of Hamilton's rule, but it would take another 10 years after 1986 to find the hidden connections. Meanwhile, throughout the 1980s and early 1990s, debate about Hamilton's rule continued.

Problems with Hamilton's rule before 1996

Hamilton's rule became a widely used standard. The simplicity of $rB - C > 0$ allowed empiricists to think through how particular natural histories might influence the evolution of phenotypes and to formulate testable hypotheses – the essential attribute for a successful theory. Most biologists continue to abide by some notion along the lines of Hamilton's rule, based on its perceived success in explaining empirical patterns (citations in Box 3).

Yet, many theoreticians vigorously attacked the simplicity of Hamilton's rule. It appeared easy to set up scenarios in which the rule failed. Those theoreticians who favoured the rule replied with ever more sophisticated theoretical analyses. Anyone with primarily biological interests, or lacking in years of specialized mathematical training, gave up following the details. Clearly, the broader notions of kin selection uniquely explained diverse aspects of natural history. Hamilton's rule seemed to capture the right idea, if not every possible assumption that one could conceive.

In this section, I discuss some of the criticisms of Hamilton's rule. I focus primarily on issues that arose before 1996. In that year, my own understanding changed with the publication of Taylor & Frank (1996). With the help of Peter Taylor's elegant insights, I came to see the proper generalization of Hamilton's rule. That generalization united the simplicity of the original rule with a new and broader scope. With the broader scope, what previously seemed like a long list of exceptions to the rule could be seen as part of an expanded way of framing problems of social evolution. Later sections discuss the changes in understanding from 1996. This section sets the necessary background by focusing on issues before that time.

Extra terms in Hamilton's rule

Hamilton's rule is not sufficient if the direction of change favoured by selection depends on some term in addition to rB and C. Queller (1985) cited several earlier examples in which an extra term is required. He then showed the general way in which such terms arise. Suppose that the phenotype, x, is the level of an altruistic behaviour that is costly to an individual, but beneficial to its neighbours. Among all the neighbours that interact within a local patch, the average level of the altruistic phenotype is x_p. Social interactions, through the effects of these behaviours, increment fitness by

$$w = Bx_p - Cx,$$

in which C is the cost to the individual for the behaviour, x, and B is the benefit received by the individual from neighbours that, on average, behave as x_p. Putting that expression into the Price equation, one finds that the altruistic behaviour increases when $rB - C > 0$, which illustrates Hamilton's rule. Here, r is the slope (regression) of x_p with respect to x, which measures how strongly an individual's behaviour is associated with the behaviour of its neighbours.

Following a variety of earlier studies, Queller (1985) pointed out that fitness might depend on synergistic

interactions between an individual and its neighbours with regard to altruistic behaviours. If the benefit only accrues when both the focal individual and its neighbour act in concert, then we need to consider a multiplicative term, $y = xx_p$. One can think of this term as describing the phenotype of pairs of individuals, in which the phenotypic value of the pair depends on how each individual in the pair behaves. For example, to achieve a task, it may be that both individuals have to contribute cooperatively to that task, otherwise no gain is achieved. In many cases, one can think of the term y as the average partnership phenotype of a focal individual paired with a randomly chosen partner from the group. If we include this synergistic effect to the increment for fitness we have

$$w = Bx_p - Cx + Dy, \tag{6}$$

in which D is the fitness contribution of the partnership phenotype. Putting this expression in the Price equation yields the condition for the altruistic behaviour to increase as $rB - C + \rho D > 0$. The term ρ is the slope of y with respect to x, which measures the association between the individual's tendency to be altruistic and the tendency of that individual's partnerships to behave in concert when faced with a task that requires joint action.

This analysis suggests that Hamilton's rule fails when there are multiplicative interactions between phenotypes, because factors beyond additive costs and benefits arise (Queller, 1985). To anticipate later discussion, note that we may think of y in eqn 6 as any characteristic other than the focal individual's value for the particular behaviour under study, x, and the average of that particular behaviour in neighbours, x_p. If ρ is the slope of y on x and $\rho \neq 0$, then the analysis will yield a condition for the increase of the altruistic character x as $rB - C + \rho D > 0$. It does not take much imagination to think of many different attributes, y, that could be associated with x and therefore cause Hamilton's rule to fail, if one chose to think of the subject in this way. Before developing that notion, let us continue with some additional issues.

Ecological context and density dependence

Hamilton (1964a,b) emphasized that limited migration would tend to keep genetically related individuals near each other. Such population viscosity could favour altruism through the increased relatedness of neighbours. A popular series of papers in the 1990s raised a problem with Hamilton's view of population viscosity (Taylor, 1992a,b; Wilson *et al.*, 1992; Queller, 1994; Frank, 1998, Section 7.1).

In a viscous population, neighbours may be related and therefore candidates for altruism. However, those same neighbours may also be the primary competitors of a potential altruist. Two potentially offsetting effects may occur. First, altruism may increase the vigour and success of a neighbour, which provides a benefit to the actor in proportion to the relatedness between actor and recipient. Second, the more vigorous neighbour may take more of the local resources, which imposes a cost on the original altruistic actor. In some cases, the two factors may cancel each other. If so, altruism cannot evolve in viscous populations, even though neighbours may be closely related.

An early analysis by Alexander

The problem was expressed beautifully in a much earlier article that is rarely cited in this context (Alexander, 1974, pp. 353, 376)

> Hamilton's development of the concept of inclusive fitness began with the argument that the reproductive success of an individual organism cannot be measured by alone considering the effects on the number and quality of direct descendants. Also involved are effects on the reproduction of genetic relatives. But, since both of these effects can only be measured in a comparative sense, there are always other individuals involved, and they are the reproductive competitors of the individuals and genetic elements being considered. In Hamilton's equations, they [the competitors] are the population at large, an average of the rest of the species. Hamilton's arguments thus seem only to consider the detriments of altruism in terms of energy expenditure and risk-taking in the act itself, and to omit or at least not specify the problem of subsequent detriment to the altruist (or its descendants) owing to the presence of the recipient (or its descendants). But all of the members of the species, or population, will not compete equally directly with any given individual. Nearby individuals are more direct competitors. This would not affect Hamilton's calculations unless nearby individuals also have a greater likelihood of being closer genetic relatives. That such a correlation generally exists is obvious, and is acknowledged by Hamilton (1972). I believe this factor modifies every consideration of whether or not, and how, nepotism will actually evolve. … The significance of this problem can scarcely be exemplified better than by a point made earlier – that if degrees of relatedness and intensity of competition among individuals diminish together in certain, not unlikely fashions with distance from any given individual in a population, then nepotism cannot evolve.

Dispersal and sex ratio

Earlier, I discussed a model of dispersal by Hamilton & May (1977). In that model, selection favoured dispersal to reduce competition against neighbouring relatives and increase competition against distant, unrelated individuals. That process of uncoupling the scales of relatedness and competition can be understood in light of Alexander's analysis.

Another line of thought independently developed the relative scales of altruism, competition and relatedness. Clark (1978) argued that 'Competition between female kin for local limiting resources may explain a male-biased secondary sex ratio …' To a mother, the benefit of making an extra daughter is offset by the competitive effect that extra daughter may have on the mother's other daughters. By contrast, sons may disperse before competing and therefore not reduce the fitness of other sons.

The development of sex ratio theory in the 1980s accounted for the interactions between dispersal, relatedness, the scale of competition and the scale of resource limitation (reviewed in Frank, 1998). That theory made clear that altruism and kin interactions could only be understood in the full ecological and demographic context of the behaviours under study. As Alexander (1974) emphasized, the scales of altruism and relatedness must be evaluated in relation to the scale of competition.

Ecology and demography in relation to Hamilton's rule

The terms r, B and C of Hamilton's rule depend on the ecological and demographic context. Any consideration of natural history makes that clear. Yet, the subject is full of 'discoveries' that Hamilton's rule fails because those terms are not constants, and that ecology and demography matter. The tension arises because simple population genetic models tend to take costs and benefits as fixed parameters rather than ecologically derived variables that depend on context.

Similarly, relatedness turns out to be part of a much broader problem of how to measure costs and benefits in common units. The traditional view of relatedness translates gene frequency deviations in an actor with respect to gene frequency deviations in a recipient, putting all terms on the common scale of consequences for gene frequency change. That makes sense. However, much confusion arose because terms that appeared to be equivalent to relatedness often popped up in analyses, yet had a variety of meanings. As we continue on through the history and the generalization of Hamilton's rule, we will see that Hamilton's rule can only be understood within a broader approach of partitioning the causes of fitness into meaningful components. Before turning to that generalization, I continue with the criticisms of Hamilton's rule that dominated discussion in the 1980s.

Further problems: dynamics

> It is better to be vaguely right than exactly wrong (Read, 1909).

The basic principles of kin selection theory and its descendant ideas always hold. Those principles are: costs and benefits of phenotypes matter; statistical associations between actors and recipients of behaviours matter; and heritability traced from the expression of phenotypes to representation among descendants matters. To most biologists, kin selection theory is understood as a concise summary of those basic principles.

The story differs in the theoretical literature. Once one loses sight of the biology, all that is left concerns mathematical aspects of the theory. Can a particular approach be used to make an exact calculation about predicted outcome? If another approach is simpler but limited in the scope over which it is exactly right, should it be discarded entirely?

These questions primarily concern mathematical rather than biological issues. Why should you care about those questions if you are interested in the biology? You should care because the theoretical literature has not done you the favour of sorting out the parts that matter to you vs. the parts that do not. Instead, each theoretical article seemingly replies to another theoretical article. The real conceptual progress that does matter remains buried under the weight of much that does not really matter to someone interested in biological problems.

Ultimately, sorting all of this out can only be done at the technical level. There is no way to argue that certain technical details do not matter without showing, technically, why they do not matter. I expressed my views about the technical issues in my book (Frank, 1998). Written 15 years ago, that book still gives a good overview of the issues. Here, I continue to avoid technical discussion, and simply evoke my perspective on the main points.

Statics vs. dynamics

> Often in the writings of economists the words 'dynamic' and 'static' are used as nothing more than synonyms for good and bad, realistic and unrealistic, simple and complex. We damn another man's theory by terming it static, and advertise our own by calling it dynamic. Examples of this are too plentiful to require citation (Samuelson, 1983).

It would be helpful to calculate exactly how selection influences phenotypes. What is the predicted sex ratio under certain patterns of mating and competition? When will individuals band together to form cooperative groups? When will groups split apart or fail because of internal conflict? The generalizations of Hamilton's rule and the broader theories of kin selection only provide exact calculations under certain assumptions (Frank, 1998).

Roughly, the theory becomes exact when the variation in fitness is small. However, significant amounts of variation are nearly universal in biology. To use kin selection as an analytical tool, when can we assume that there is little variation? Lack of variation primarily arises when a population approaches an equilibrium.

Near an equilibrium, little further change occurs, and the population comes nearly to rest – the condition of stasis. Analysis near an equilibrium is sometimes called 'statics'. As an exact analytical tool, kin selection theory is primarily a tool for statics.

The problems of statics are widely known. Real biological systems are unlikely to be near an equilibrium. Thus, exact analysis must consider dynamics – the full processes of change. Worse, many problems have different points at which the system could come to rest – alternative equilibria. If one only analyses what happens near an equilibrium, as in statics, one has no idea which of the alternative equilibria the system will end up near. Only a full dynamical analysis of change can indicate which of the equilibria one would expect the system to evolve towards.

Given all of the benefits of dynamics and the limitations of statics, why would anyone ever consider a theory based on statics? Because, to make a dynamic analysis, one has to make a lot of exact and very specific assumptions, otherwise one cannot do the analysis. For example, one usually has to specify exactly how the genetics of a phenotype is controlled in order to make a complete model of population genetics. The problem is that we do not know the proper assumptions to fill out the required list for a dynamic analysis. So, in making all of the necessary detailed assumptions for a dynamic analysis, one is left with an exact calculation that applies exactly to nothing.

By contrast, the static analysis requires few assumptions. In kin selection arguments, usually one needs to specify how phenotypes and various environmental factors influence fitness. One obtains a static analysis that translates the biological assumptions about fitness into a prediction about how natural selection influences the evolution of phenotypes. Clearly, this is a vague sort of analysis, but it is not exactly wrong, as is the full dynamical analysis. Put another way, a static analysis does not suffer the pretense of exactness. Instead, a static analysis accepts the limitations and calculates the qualitative predictions about what one expects to see in various natural settings.

Comparative statics

> Now, an observer fresh from Mars might excusably think that the human mind, inspired by experience, would start analysis with the relatively concrete and then, as more subtle relations reveal themselves, proceed to the relatively abstract, that is to say, to start from dynamic relations and then proceed to the working out of the static ones. *But this has not been so in any field of scientific endeavor whatsoever:* always static theory has historically preceded dynamic theory and the reasons for this seem to be as obvious as they are sound – static theory is much simpler to work out; its propositions are easier to prove; and it seems closer to (logical) essentials (Schumpeter, 1954).

A static analysis summarizes the major forces that potentially influence phenotype. However, that sort of simplified theory cannot predict the actual phenotypic value that one expects to observe. Instead, one should think of statics in terms of comparison. The dispersal model discussed earlier provides a good example. In that model, the predicted dispersal probability from a static kin selection analysis was given in eqn 5, repeated here

$$d^* = \frac{r - c}{r - c^2},$$

in which r is the relatedness between competitors on a patch, c is the cost of dispersing, and d^* is the predicted equilibrium dispersal rate. Certainly, no one believes that this model will predict the actual dispersal rate in real cases. Too many factors are left out. Instead, the whole idea of the analysis is to isolate a few key processes.

The relatedness term, r, raises another problem. In an actual population, the relatedness in a patch would depend on the dispersal rate. The theory given above does not tell us how to analyse this dependence between relatedness and dispersal. To simplify the analysis, we could set relatedness, r, as a given parameter. The model then predicts that as the relatedness among competitors on a patch increases, the observed dispersal rate increases. That sort of prediction is the method of comparative statics.

Comparative statics begins with the assumption that a dynamic analysis following the joint dependence between dispersal and relatedness would, ideally, be preferable. However, one cannot easily achieve that ideal, because dynamical analysis requires specific assumptions about a variety of processes for which we do not have information. Instead, one admits that one does not know enough to predict dynamics, and so the analysis should emphasize statics. A static analysis is based on fewer, simpler assumptions.

The comparative static analysis isolates key causal processes in a direct way. For example, the dispersal model makes an interesting comparative prediction: as the relatedness in a patch increases, the predicted dispersal rate increases. The ideal empirical test identifies natural or laboratory settings in which relatedness changes and one can measure the associated change in the amount of dispersal. The goal is to measure the change in the putative cause and the change in the outcome. If the direction of change in the outcome repeatedly tends to follow the predicted direction of change, then one is on to something.

A practical method for solving problems and recovery of Hamilton's rule

Hamilton did not use kin selection theory to analyse models of phenotypes. He did pass on to his students

unpublished notes about how the Price equation might be used to study sex ratios. I obtained those notes in 1979 while attending Hamilton's graduate seminar. I modified Hamilton's Price equation method to solve a wide variety of problems with kin interactions, including the dispersal model that I presented earlier. Hamilton's class notes are posted as Supporting Information for this article in the files Appendix S1 and Appendix S2.

The Price equation method was a bit tedious. Eventually, I found the match between the standard ESS analysis of phenotypes and the Price equation analysis of kin interactions. Earlier, I briefly mentioned the enhanced ESS approach for kin interactions. This section describes that enhanced approach in more detail and considers the broader consequences for understanding the theory.

First steps

The Price equation is a general expression for the change in average phenotype. To calculate the change in phenotype, one needs to specify the relation between phenotype and fitness. For example, Box 4 shows the relation between dispersal and fitness. When one puts that expression for fitness into the Price equation, one obtains a variety of terms. Each term describes the contribution of a component of fitness to the overall change in phenotype (Frank, 1986a).

In a typical problem with kin interactions, different components of fitness oppose each other. Some favour the increase in the phenotype, others favour the decrease in the phenotype. For example, dispersal imposes a direct cost on an individual, because the increased risk during dispersal increases the chance of death before reproduction. By contrast, dispersal benefits the reproduction of neighbours by reducing the competition experienced by those neighbours. Together, the various terms in the Price equation analysis define the different components of fitness.

I found that I could avoid using the Price equation by directly calculating the value of the phenotype that maximized fitness. The maximization approach came from analyses of ESS phenotypes (Maynard Smith, 1982). The idea is simply that, at equilibrium with regard to selection, the favoured phenotype must have fitness at least as great as any slightly different phenotype. If that were not the case, then the nearby phenotypes with higher fitness would increase, and the previous candidate phenotype was in fact not an equilibrium with respect to nearby alternatives.

To find a local maximum, one uses the standard calculus approach. Write a function that relates phenotypes to fitness. Take the derivative (change) of fitness with respect to phenotype. That derivative describes whether fitness increases or decreases with respect to a change in phenotype. One assumes that all members of the population have the same phenotype, and then analyses the fitness change for a small fraction of the population that has a slightly deviant phenotype. In other words, little variation exists.

A local maximum can only occur when the fitness neither increases nor decreases for the deviant phenotype. If the derivative were increasing, then larger phenotypes would be favoured. If the derivative were decreasing, then smaller phenotypes would be favoured. When the derivative is zero, then a balance in forces has been achieved, and no change is favoured. When at a balance, one checks larger phenotypic deviations to make certain that fitness would indeed decrease if the phenotypic value changed. If so, then the balance point, where the derivative is zero, is a local maximum and an ESS. Maximization is just a mathematical trick for finding an equilibrium. An equilibrium is a point at which stasis occurs, leading to a static analysis.

When applying the maximization method, one ends up with terms that are the change in partner phenotype with respect to the change in the focal individual's phenotype. As I mentioned earlier, that kind of term was initially thought to be difficult to interpret when kin interactions occur. When partners are genetically related, then the association between partner phenotype and focal individual phenotype depends on the degree of genetic similarity. Early analysts recognized that relatedness matters (Hamilton, 1967; Hamilton & May, 1977), but abandoned the method because it was not clear exactly how genetic similarity would translate into the relation between partner phenotype and focal individual phenotype that arises in the calculus method of differentiation.

I matched the Price equation terms to the ESS maximization method. I could see that the difficult term for the change in partner phenotype with respect to individual phenotype was like the regression term that Hamilton (1970) had come to use as the coefficient of relatedness in his theory. That equivalence meant that one could use the much simpler maximization trick and the power of calculus to analyse fitness and find ESS phenotypes. One just had to replace the change in partner phenotype with respect to focal individual phenotype by the coefficient of relatedness between them.

Before 1995, I only published Price equation analyses. That method, although a bit tedious, easily gave solutions to many problems of dispersal, sex ratio and tragedy of the commons models for sociality. I published a series of articles between 1985 and 1994 on those topics, as summarized in Frank (1998). I did not publish the maximization method without use of the Price equation, because the Price equation had a prior history in the literature and seemed like a more defensible approach. However, in Frank (1995), I analysed a more complex problem, for which avoiding the Price equation and using only the calculus method provided

important advantages in understanding the evolution of phenotypes in social interactions. Thus, I needed to develop the calculus approach into a publishable form.

Generalized Hamilton's rule as a marginal value expression

To develop the calculus method, I approached Peter Taylor in 1995. Taylor found a way to connect the calculus method to Hamilton's rule. I had abandoned Hamilton's rule, because nothing like Hamilton's rule had appeared in my numerous studies of different phenotypes. Suddenly, all of my prior analyses could be understood more deeply by their connection to the generalized Hamilton's rule that came from our work (Taylor & Frank, 1996).

The calculus method automatically separates out the various causes of fitness into the three aspects of Hamilton's rule. First, all of the focal individual's phenotypic effects on its own fitness combine into one term. In the calculus analysis, that term is the small change in direct fitness for a given small change in the focal individual's phenotype, holding constant the phenotype of other individuals.

The small changes are usually described as 'marginal changes,' matching the classical usage of marginal values in economic analysis. The first term is thus the marginal effect of an individual's phenotype on its own fitness, holding constant all other effects. This marginal effect matches exactly the cost term in Hamilton's rule. One traditionally defines this effect as a cost in such models, because in the standard case of altruism, one analyses a phenotype that directly lowers the actor's fitness – the cost. Mathematically, there is no need for this direct effect to be negative and costly, and in some models it is not. However, we retain the traditional usage and label this term 'the cost'. Because the calculus approach analyses small changes, that method automatically gives us the marginal cost.

The second component is the marginal effect of small changes in an individual's phenotype on the fitness of social partners. Traditionally, the effect of the actor on recipients is called the 'benefit'. Thus, this second effect is the marginal benefit component.

The third component measures the association between the actor's phenotype and the phenotype or genotype of social partners. The exact measure depends on various issues (Frank, 1998). Here, simply note that the calculus approach automatically weights any marginal benefit components by the association between the actor and recipient. That association matches the coefficient of relatedness from Hamilton's rule, although in a generalized form.

An equilibrium can occur only when selection does not favour a change in phenotype. Thus, at equilibrium, the marginal Hamilton's rule equals zero and has the form

$$rB_m - C_m = 0, \tag{7}$$

in which I use the m subscripts to emphasize that the benefit and cost terms are marginal values (Taylor & Frank, 1996). The marginal values will change with changing phenotypic values and with changing ecological and demographic context. Thus, this analysis makes clear that Hamilton's rule arises from context-dependent benefit and cost terms. Others had noted the context dependence of those terms (Grafen, 1985; Queller, 1992a,b). However, Taylor & Frank (1996) was the first approach that easily found the ESS phenotype and at the same time showed the underlying conceptual basis by expressing a generalized Hamilton's rule. I use the term 'generalized' because the analysis extended the kinds and complexity of social interaction that could be studied and the interpretation of relatedness. It also became clear how to deal with multiple social processes simultaneously, leading to multiple marginal cost and benefit terms.

The Taylor & Frank method leaves out much mathematical detail. That simplicity allows one to start with an expression for how different phenotypes and other factors influence fitness, take a standard type of derivative from calculus, and evaluate an equilibrium ESS outcome favoured by selection. This method of analysis automatically gives the form for the equilibrium ESS condition as the marginal Hamilton's rule in eqn 7. The ESS provides the basis for comparative statics, in which one can see clearly how the predicted phenotype changes with respect to various social, ecological and demographic causes.

To achieve that simplicity, one ignores dynamics, certain details of genetics and the developmental complexities that connect genotype to phenotype. Most of those complications do not alter the mathematical results with respect to searching for equilibrium values. That magical simplification arises because, whenever the amount of variation in the population is small, most of the complexities become negligible in size and the method correctly identifies the outcome of selection. Although it is possible to include additional complexity, the method simplifies by design, following the precepts of comparative statics.

Marginal Hamilton's rule analysis of dispersal

The marginal Hamilton's rule in eqn 7 evaluates a problem by the marginal costs and benefits of a phenotype. Consider the dispersal model of Hamilton & May (1977) described earlier. If we start with the fitness expression in Box 4, take the derivative of fitness with respect to the variant phenotype of a focal individual, and collect together the terms, we end up with the cost, benefit and relatedness components of the marginal Hamilton's rule (Frank, 1998, Section 7.2).

Following that method

$$C_m = \frac{c}{1 - cd}, \tag{8}$$

in which the marginal cost of increased dispersal, C_m, is the cost of dispersal, c, divided by the level of competition on a patch for a breeding spot, $1 - cd$ (Box 5). The marginal benefit is

$$B_m = \frac{1 - d}{(1 - cd)^2}, \tag{9}$$

in which the numerator, $1 - d$, describes how increased dispersal reduces the competition experienced by neighbours. The denominator adjusts the benefit of reduced competition by the intensity of competition between pairs of individuals for each breeding spot $(1 - cd)^2$ (Box 5). Putting these terms in the marginal Hamilton's rule of eqn 7 leads to the general solution for dispersal in eqn 5.

The marginal cost and benefit expressions are essentially impossible to obtain either intuitively, by thinking about the dispersal problem, or by inspecting the mathematical expression for fitness given in Box 4. It is only with the maximization technique that the separate marginal cost and benefit expressions can be found.

Once one has the marginal expressions, one can study them to learn how the simple biological assumptions translate into cost and benefit effects on different components of fitness. As the phenotype of interest changes, the costs and benefits change through complex social, demographic and ecological interactions. There will essentially never be fixed costs and benefits that can be plugged into some equation. Instead, one must extract those costs and benefits from the biological assumptions and the analysis of the problem.

Historically, people have tended to take Hamilton's rule as an expression based on fixed costs and benefits. That history arose because of the population genetic modelling that was associated with the early evaluation of the theory. In a population genetic model, the tendency has always been to set costs and benefits as parameters associated with different genotypes. Frequency and density dependence, and other biological interactions, were considered distinct from the costs and and benefits, as if those costs and benefits were not an outcome of the biology. That approach misled people to think of Hamilton's rule as an expression that translated fixed costs and benefits into a conclusion about evolutionary change.

The Taylor & Frank (1996) method helps because it gives a way to translate biological assumptions about phenotypes and fitness into the separate marginal cost and benefit terms needed to evaluate Hamilton's rule. One must keep in mind that Hamilton's rule is, in practice, an expression that derives particular meaning from context. Without context, the marginal cost and benefit terms are abstract. That abstraction is the primary strength of Hamilton's rule, allowing it to apply universally. At the same time, abstraction often causes confusion, because the power of abstraction requires a certain degree of consideration to understand and apply properly (Frank, 2012b).

Box 5: Analysis of dispersal model

For the marginal cost in eqn 8, we obtain the competition term in the denominator by counting the number of individuals on a patch competing for each breeding slot. That number is proportional to the fraction of individuals that do not disperse and stay at home to compete, $1 - d$, plus the fraction of individuals that disperse and become immigrants into each patch. The immigrant fraction is the fraction that disperse, d, multiplied by the probability of surviving the dispersal phase, $1 - c$. Putting the pieces together for the denominator, which is an expression proportional to the number of competitors on a patch, we obtain $1 - d + d(1 - c) = 1 - cd$.

For the marginal benefit in eqn 9, the denominator measures the intensity of competition between pairs of individuals in the patch for a given level of dispersal and cost of dispersal. Roughly speaking, intensity of competition can be measured by the probability that two individuals will compete for the same breeding spot. From the previous paragraph, the intensity of competition for a breeding spot is proportional to $1 - cd$. Thus, the pairwise competition for a spot is proportional to $(1 - cd)^2$.

Taylor's insight into reproductive value, demography and life history

In the recent literature, the methods of Taylor & Frank (1996) are primarily used to analyse specific models by the marginal version of Hamilton's rule (eqn 7). As long as one accepts the approach of comparative statics, generating testable hypotheses follows a simple procedure. First, express the different phenotypes and ecological factors that affect fitness in an equation that describes the natural history assumptions. Second, maximize the fitness expression to obtain the predictions of comparative statics.

That mathematical method solves what had previously been complex and often beyond analysis. It also provided a conceptual advance by developing my earlier Price equation and maximization approaches into the marginal Hamilton's rule expression in eqn 7. The marginal Hamilton's rule clarified understanding about how selection works and how to interpret a wide variety of problems in natural history (Frank, 1998).

The importance of reproductive value

From my point of view, however, the great advance in Taylor & Frank (1996) came from Peter Taylor's insight about reproductive value in models of kin selection.

Reproductive value concerns the relative contribution of an individual to the future of the population. For example, older individuals may have a lower expectation of future reproduction before death than do younger individuals. The benefit of altruism provided to an older individual must be discounted by the low expected reproductive value of old age when compared with the greater reproductive value associated with altruistic benefits given to a younger individual.

The problem of reproductive value had always been a part of kin selection theory. Hamilton (1972) clearly noted that different individuals in a social interaction may contribute differently to the future of the population. In Hamilton's analysis, when calculating the benefit of altruism towards different individuals, one must weight each individual by its genetic relatedness to the actor and by its relative reproductive value. For example, helping a young cousin in the prime of life provides greater net benefit than the same help given to a much older sibling near the end of life. Relatedness alone is not sufficient.

Prior status of the theory

Although the importance of reproductive value was understood, the theory was in an odd state before 1996. The general approach used by Hamilton and most followers was simply to attach a reproductive value weighting to each component of fitness. If an actor and recipient had different reproductive values, then an extended Hamilton's rule might be $rBv_r - Cv_a > 0$, in which v_r and v_a are the reproductive value weightings for recipient and actor. If, for example, the recipient is a young individual near prime reproductive age, and the actor is an old individual near the end of life, then v_r is much larger than v_a, and the old individual may be favoured to express a costly altruistic act towards the young individual even if the recipient is only distantly related to the actor.

Simply, attaching reproductive value terms to fitness components is correct but often not helpful. It is not helpful, because it provides no guidance about how to find the right valuations in realistic problems.

Separately, a complete theory of life history had developed (Charlesworth, 1994). That theory provided clear guidelines for how to compare different classes of individuals and different components of fitness with respect to their reproductive values. The relative reproductive value weightings predict how life history characters, such as relative investment in reproduction and survival, may change with age, condition, local ecological factors and the demographic structure of the population. That well developed theory of life history had many successes in explaining major aspects of organismal physiology and behaviour.

Actual problems of natural history often required combining the relative reproductive valuations with the role of kin interactions. However, connecting life history theory to kin selection analysis had not been achieved in a generally useful way. When studying the theory of such problems, one could come up with various special approaches or complicated analyses (e.g., Frank, 1987). However, if one cannot apply a theory in a clear and simple way, one does not truly understand the theory at a deep level. In this case, the basic issue was combining reproductive value with the various aspects of phenotypic evolution that arose in social settings. But it was not known how to do that in an easy way.

Combining kin selection and life history

Peter Taylor saw all of that. He figured out how to attach our general kin selection approach with life history analysis. The extended method accounted for the full context of ecological and demographic factors that interact with social phenotypes (Taylor & Frank, 1996). The method easily translates natural history into analysis. As before, one specified a problem by how phenotypes affect fitness. In addition, one now also specified the different classes of individuals involved and the distinct components of fitness. For example, there could be older individuals and younger individuals. The age groups would be embedded in the demographic structure of birth and death rates, which could depend on the social phenotypes. Each class could be either actor or recipient or both, in the sense that the phenotype of interest could be expressed by different kinds of individuals and could affect different kinds of individuals. Many realistic problems of sociality have this sort of class structure.

All of this may sound complicated. But the beauty of Taylor's insight is that the crude maximization method that I had been using was transformed into a simple method that combined analysis of sociality, generalized notions of kin relations through correlations between phenotypes, and the full power of life history analysis.

Opportunity for synthesis

Up to 1996, I had relatively little interest in theories of kin selection, inclusive fitness and group selection as separate subjects worthy of study. Instead, I had thought of those theories as tools that one used to solve problems of natural history, in the sense of comparative statics. I had developed those solutions into testable hypotheses about interesting phenotypes. My collaboration with Peter Taylor opened my eyes to the deep problems of selection that had been latent in the subject.

In particular, our new approach brought out the proper formulations for marginal valuation and reproductive valuation. Of course, that was, in a sense, not

new, because Hamilton had all the right ideas. But, just as Hamilton could not use kin selection theory to solve the problems that most interested him, such as dispersal and sex ratio, he also could not use life history and reproductive value to solve problems of sociality embedded in the natural complexities of demography and multiple classes of social interactants.

I had always believed that if one could not use a theory to solve problems, then one had only a superficial understanding of deeper concepts. With Taylor & Frank (1996), I suddenly realized the deep connections between the understanding of marginal valuation and reproductive valuation in sociality and the solving of actual problems of natural history. That new perspective caused me to take up the study of the general theory. I wanted to evaluate the status of the deeper principles in relation to the way in which one analysed problems.

Unresolved issues

I soon found that Taylor & Frank (1996) led to many key points that remained vague or unresolved. Two issues stood out. First, the meaning of the relatedness coefficient seemed to be confused. Variations in interpretation arose for different kinds of problems. The original simplicity of Hamilton's genetic kinship theory had finally sunk. There were many earlier hints of problems, but with the expanded scope of the theory, the contradictions became too numerous to ignore.

The second issue concerned the connections between the models of phenotype in sociality and the developing theories of multivariate selection in quantitative genetics. Those theories of multivariate selection had advanced greatly since the classic article by Lande & Arnold (1983). Following Lande & Arnold, many studies considered how to partition the causes of selection among the different characters that affected fitness and how to analyse expanded notions of heritability. Those advances in quantitative genetics seemed to be closely related to the problems of analysing social evolution. Some steps had been made to connect those advances to sociality, yet the deeper relations remained unclear.

Following up on what I learned by working with Peter Taylor, I set out to understand those two issues: the generalization of relatedness and the causal analysis of fitness and heritability. I eventually came across a surprising number of problems that I had never understood or had not even realized that I was ignoring in my many prior analyses of kin interactions.

Queller's insight and the true meaning of Hamilton's rule

In pursuit of unresolved issues, I soon came to the key articles by Queller (1992a,b). Queller developed the idea that Hamilton's analysis had always been about the causal interpretation of fitness components. Hamilton separated the total fitness effect of a social act into costs, benefits and relatedness so that one could reason more clearly about how selection shaped behaviours. Hamilton never presented the theory as an alternative to classical methods. Instead, he was after causal decomposition.

Much of the literature through the 1980s lost sight of this primary emphasis on causal decomposition. Controversies about population genetics vs. kin selection were ultimately about the tension between dynamics and comparative statics. Full dynamical analyses with detailed assumptions about genetics provide exact theories that perhaps apply to no real cases. Statics applies approximately to all situations, but perhaps not exactly to any particular case.

In my own comparative statics analyses of dispersal, sex ratios and various social traits, I focused on the solutions for phenotypes in terms of biological assumptions. I did not try to analyse the problems with respect to the sort of causal decomposition into costs and benefits that Hamilton had emphasized. Later, I came to understand the power of the marginal Hamilton's rule. I then began to understand my past comparative statics models in terms of Hamilton's partition into causal components. In particular, the marginal Hamilton's rule automatically separated fitness components into direct effects (costs) and indirect effects (benefits) weighted by relatedness (Taylor & Frank, 1996).

Queller (1992a,b) had also derived an expression similar to the marginal Hamilton's rule. Queller first partitioned fitness into components according to a regression model. His regression terms included the phenotype of the actor and the phenotype of the recipient – or the genotypes depending on the analysis. Queller considered the different phenotypes within the general context of quantitative genetic analyses of multivariate selection (Lande & Arnold, 1983). That connection to multivariate quantitative genetics eventually opened the way to a broader interpretation of the components of fitness and the components of heritability.

Queller put his multivariate regression expression for fitness into the Price equation to obtain a multiple regression form of Hamilton's rule. One term is the effect of an actor's phenotype on its own fitness, holding constant the phenotype of its neighbours (cost). The other term is the effect of the neighbours' phenotype on the actor's fitness, holding constant the actor's phenotype (benefit). These cost and benefit terms arise as partial regression coefficients in the partitioning of fitness into components. In the Price equation, one automatically obtains a weighting of the benefit term by the regression of neighbour phenotype on actor phenotype (or genotype). That weighting provides a measure of relatedness. Thus, the condition for the increase of an altruistic behaviour is $rB - C > 0$, where each term directly corresponds to a regression coefficient.

Queller's partial regression terms for cost and benefit are similar to the cost and benefit terms of the marginal Hamilton's rule. However, two differences are important. First, Queller's approach emphasizes the causal analysis of components. He explicitly related the regression model of fitness to a path analysis model, which highlights a causal interpretation of the different terms in the regression.

Second, the regression terms, although more general and potentially interpreted by causal analysis, provide a poor method for analysis of actual problems in terms of comparative statics. Even simple models of dispersal and sex ratio lead to complex and essentially uninterpretable regression expressions. That complexity led me to abandon the direct use of such regressions in my earlier Price equation models and instead to develop the maximization technique. The maximization technique requires the additional assumption of limited variation. In return for that assumption, one obtains a simple and powerful comparative statics tool. In practice, one trades the conceptually powerful regression modelling and causal analysis for the analytically powerful maximization method and comparative statics.

Once I had the first formal expression of the maximization technique from Taylor & Frank (1996), I evaluated the advantages and disadvantages of the various approaches. I could see that one had to unite Queller's causal approach through regression and path analysis with the analytical power of the maximization and marginal value techniques. Causal analysis ties back to Hamilton's original goal of separating fitness and transmission into components to reason more clearly about how selection and genetics shape social traits. Marginal values and comparative statics are also necessary, to provide the tools to analyse actual problems. Put another way, causal analysis provides the foundations for reasoning about complex problems, and marginal value analysis provides the techniques for applying that reasoning to particular cases.

Taylor & Frank (1996) came close to uniting the quantitative genetic and causal models given by Queller (1992a,b) with the calculus techniques for comparative statics. However, once that loose connection was made, I soon began to see the next level of unsolved problems. In essence, the overly simple notions of an actor and a recipient and of a cost and benefit were too limiting. Those restrictive assumptions limited general understanding of causes and the analysis of particular cases. To move ahead, one had to think through various types of natural history and to work out how to separate the many causes into distinct components. That causal decomposition was Hamilton's original goal. But one no longer had to be confined to the oversimplified abstractions in which Hamilton originally worked to show how causal decomposition might be done. Instead, the time had come for a more general approach.

The class structured modelling for social interactions introduced in Taylor & Frank (1996) suggested the next steps. Taylor & Frank gave examples with multiple classes in social interactions. Those sorts of multiclass problems lead to more general causal decompositions with multiple actors and recipients and, more generally, with a broader way to reason about the causal components of realistic problems. Those multiclass problems also brought out the challenge of tracing the distinct pathways of transmission and heritability that determine what fraction of the change by selection transmits to the future population.

The way forward was to work carefully with causal decompositions of fitness and the transmission of characters, to use the Price equation as a formal tool to keep track of everything in an exact evolutionary analysis, and then to simplify as needed to obtain practical tools for comparative statics. Although that may sound complicated, it turns out to be a natural extension of $rB - C > 0$. One simply needs additional causal terms to match realistic problems, and to interpret the various terms more broadly than Hamilton did.

The proper generalization of kin selection theory

Hamilton's theory is ultimately about causal interpretation. The proper generalization arises from a clearer understanding of causal decomposition. With regard to the causal analysis of relatedness, Hamilton (1970, 1975) had already given up on the limited interpretation of the theory with regard to pedigree relations and classic notions of kinship. Instead, he realized that the theory could be extended to deal with genetic similarities no matter how those similarities arise. Causally, selection must be indifferent to the process that generates genetic similarity. Selection can only act on the current patterns of genetic variation and the current processes that influence fitness.

Subsequent generalizations continued to refine the causal interpretation of selection. The theory naturally transformed from its initial emphasis on identity by descent and lineal kin relations to statistical associations of genotypes and then to broader aspects of correlated characters in social interactions.

An example: interspecific altruism

My study of altruism between species taught me that kin selection must be thought of as part of a wider set of problems (Frank, 1994). I had asked the typical altruism question: when is an individual favoured to help another at a cost to itself? In this case, the problem concerned whether an individual of one species could be favoured to help an individual of another species. Clearly, the traditional view of genetic kinship could not be involved. Members of different species

that do not interbreed cannot be cousins or other types of related kin.

I set up the interaction between species as a variant of Hamilton's standard model. In this case, I evaluated whether altruistic behaviour by one species toward a second species can increase by selection.

Individuals of the first species have phenotype x, the level of help they provide to individuals of the other species. The altruistic phenotype directly reduces the fitness of actors from the first species by Cx. Individuals of the other species have phenotype y, the level of help they provide to individuals of the first species. That altruistic phenotype of the second species directly enhances the fitness of recipients of the first species by By. The goal is to evaluate how these behavioural interactions determine the direction of change in the altruistic phenotype x of the first species.

The first species cannot accrue inclusive fitness benefits by helping the second species. An inclusive fitness benefit is the indirect transmission of the actor's genotype through the recipient of the actor's altruistic behaviour. Members of another species cannot carry genotypic differences that influence the evolution of traits in the focal species. Inclusive fitness has no meaning in relation to altruism between species. Nonetheless, we end up with Hamilton's rule, as follows.

Focus on an individual in the first species with altruistic phenotype, x. That altruism reduces fitness by Cx. The focal individual may receive benefits from the altruism of the other species. Suppose that the particular social partners from the other species for the focal individual have phenotype y, which adds By to the fitness of the focal individual. The combination of gains and losses for these effects causes an increase in fitness to the focal individual when

$$By - Cx > 0.$$

Now comes the key step: the association between the altruistic behaviours of partners from the two species. Suppose the slope (regression) of y relative to x is r. Then, in evaluating fitness changes, we can use $y = rx$, because the altruistic level of the focal individual, x, predicts the associated value of y. The translation between x and y is the regression coefficient, r.

Using $y = rx$, the fitness change is positive when

$$Brx - Cx > 0,$$

which is the same as

$$rB - C > 0.$$

Is that Hamilton's rule? If one abides by inclusive fitness and the traditional view established by Hamilton, then the answer is no. If one recognizes that the traditional view came into being before we understood the broader analysis of characters and the general role of correlations, then the answer is maybe. In the latter case, we must figure out the broader context and its relations to traditional models of social evolution, and then make a decision about how to understand the full range of social characters.

After Taylor & Frank (1996), I followed up by trying to apply the new theory to various problems of natural history. Several conceptual limitations became clear. The most important problem concerned the meaning of relatedness. A second associated issue concerned the interpretation of inclusive fitness. The remainder of this section discusses relatedness. The following section takes up inclusive fitness.

Two types of relatedness: social partners and transmission

Queller (1992a,b) quantitative genetic approach linked kin selection to Lande & Arnold's (1983) general analysis of multivariate selection. In the traditional multivariate approach, one usually thinks of two different phenotypes as traits present in each individual. That same conceptual approach to multivariate analysis can also evaluate social situations, in which one phenotype is present in one type of individual, and the other phenotype is present in the social partners of the first type (Frank, 1997b).

The two interacting types of individuals may be players in a game, members of different species, members of a family or any other combination. For any of those interactions, we can often evaluate the consequences in the same way that I described for the interaction between two different species. The focal individual has a phenotype with direct cost Cx. That focal individual also has social partners that influence the focal individual's fitness by a factor By, where y is the average phenotype of the social partners. These two phenotypes, x and y, lead to a multivariate analysis of selection that depends on the correlation between characters. In this case, the two characters happen to be in different individuals, but the analysis is essentially the same as a standard multivariate selection problem.

The correlation between the focal individual's phenotype, x, and its partners' phenotype, y, is best expressed as a regression coefficient r of y on x. If partners are genetic kin, the r will be a kind of genetic kinship coefficient. If the partners are genetically similar but not by traditional lines of kinship, the r is still similar to the form of the relatedness coefficient that Hamilton (1970) introduced as the key to his theory, which did not depend on traditional kinship ties.

The partners may have correlated phenotypes, but be genetically uncorrelated. If so, our focal individual still gains by the beneficial social phenotype of its partners. The magnitude of that gain is in proportion to the same regression coefficient, r, in which the association is purely phenotypic. However, we must separate two aspects: selection and transmission. That separation forms a standard part of multivariate quantitative

genetics. Selection is the differential success within a period, such as a behavioural episode or a generation. Transmission is the fidelity by which selected traits are transmitted to the future, the heritability.

In this model, we can think of two distinct classes of individuals. The focal individual is both an actor and a recipient for its own phenotype, x. It is an actor because it expresses the phenotype; it is a recipient because its success is influenced by the same phenotype, x. The partners are actors, because they express the phenotype, y. But they are not recipients, because we have not specified that either trait, x or y, affects the partners' own success. The focal individual is a recipient of the phenotype, y.

The total fitness increment on the focal individual with respect to the phenotype, x, is proportional to $rB - C$, as shown in the previous section. Evolutionary change depends on the heritability of the phenotype, x. This follows the typical combination in genetics: selection determines relative reproduction, and heritability determines the fraction of selective change that is transmitted to the future.

The heritability is not particularly important in this case. Suppose, for example, that the heritability is $\tau = Rh$. Here, R is the genetic kinship between the focal individual and its descendants, and h is the fraction of the phenotypic variability attributed to genes. In a typical parent–offspring example, $R = 1/2$. If we compared an individual to a niece through a full sibling, then typically $R = 1/4$. The condition for x to increase must include the heritability. Thus, the condition is

$$(rB - C)\tau > 0,$$

which describes the transmitted fraction, $\tau = Rh$, of the selective gain, $rB - C$. If the trait is heritable, $\tau > 0$, then this condition is the same as $rB - C > 0$. Here, distinguishing the similarity between social partners, r, and transmission in proportion to R, did not change the result.

Distinguishing the types of relatedness

In many cases, one must distinguish the role of social partners from the role of transmission. Consider the following example (Frank, 1997b, fig. 9). There are two classes of individuals. Individuals of the active class express an altruistic phenotype, x, with cost C. A focal individual of the active class has social partners from the active class that express an average level of altruism, y, with benefit B to the focal individual. The fitness increment for a focal individual in this active class with respect to the altruistic behaviours is $w_1 = By - Cx$.

Each individual from the active class also has a single partner from a second, inactive class. Members of that second class do not express an altruistic phenotype. They may, for example, be relatives that receive care but do not give care. The inactive partner gains a benefit, $\hat{B}$, from its active partner's altruism, x, so its fitness increment from altruism is $w_2 = \hat{B}x$. To simplify, we ignore the potentially different reproductive values of the two classes. An inactive partner might, for example, be an offspring or a nondescendant kin.

From the previous section, the direct fitness effect on class one is $rB - C$, where r is the regression of y on x. If the phenotypic association between an actor and its social partners, r, arises from genetic similarity, then r is a classic kin selection coefficient of relatedness. However, nothing in the model requires genetic similarity. Here, r only has to do with phenotypic similarity, because $rB - C$ is the fitness effect separated from aspects of transmission to the future.

The class one individuals transmit their phenotypes to the future in proportion to τ_1, the heritability of their altruistic trait. Thus, the total direct transmitted component by class one is proportional to $(rB - C)\tau_1$, which matches the result of the prior section.

In this case, we also have the beneficial effect, $\hat{B}$, of the class one individuals on their inactive partners from class two. The benefit, $\hat{B}$, describes the fitness effect separate from aspects of transmission to the future. For example, if the class two individuals are genetically unrelated to their actor partners, then the enhanced fitness of the class two individuals has no effect on the evolution of altruism, because those class two recipients are genetically unrelated to the actors.

In general, we may specify the heritability of a class one actor's phenotype, x, through its beneficial effect on its class two partner, as τ_2. For example, the class two partner may produce nieces and nephews of the actor, and τ_2 would be the relatedness of the actor to its nieces and nephews. However, the more general interpretation is important: τ_2 is the heritability, or transmitted information, of the class one phenotype through its beneficial fitness effect on its partner of class two. The net direct effect of selection and transmission on class two is $\hat{B}\tau_2$.

Combining the class one and class two effects, the total transmitted consequence of the actor's phenotype favours an increase in altruism when

$$(rB - C)\tau_1 + \hat{B}\tau_2 > 0. \qquad (10)$$

This expression combines the role of correlated phenotypes between social partners, r, and the pathways for the fidelity of transmission to the future, τ.

This example illustrates how to combine the two different aspects of similarity, or relatedness, that arise in models of social evolution. The general approach requires separating classes of individuals according to their role in the social process, following the direct fitness effects on each class, weighting each class by the fidelity of transmission of the phenotype under study and also weighting each class by its reproductive value

(Frank, 1997a,b, 1998). In practice, one typically uses the maximization technique of Taylor & Frank (1996), as updated in Frank (1997a, 1998).

All of this follows the kind of causal decomposition at the heart of Hamilton's approach to kin selection theory. However, one has to accept several generalizations to the theory, otherwise the problem is beyond understanding by kin selection analysis. In particular, one must separate the correlation between phenotypes that influences fitness from the correlation between an actor and various descendants that determines heritability.

The original theories of kin selection and inclusive fitness blurred the distinction between these different kinds of correlation. To understand the issues more broadly, one must accept a theory that follows different causes of fitness, including correlations between social partners, and different causes of transmission, including direct and indirect pathways by which phenotypes pass to future generations (Frank, 1997a,b, 1998).

Kin selection vs. correlated selection

I have emphasized a causal analysis of selection rather than a purely kinship analysis of selection. In the broader causal perspective, the two key factors are transmission and selection. Transmission of social characters always depends on aspects of shared genotype, or at least on shared heritable traits. For selection, correlated social phenotypes play the key role. Such correlations may arise by kinship, by shared genotype through processes other than kinship, or by associations through processes other than shared genotype.

With regard to correlated social phenotypes, it may seem tempting to define the genetic associations as the proper limited domain for kin selection theory. Hamilton developed his theory by first analysing classical pedigree kinship. He then broadened his scope to include shared genotype through processes other than kinship. But he did not expand his theory to the general analysis of correlated traits by processes other than shared genotype.

There can never be a final resolution with regard to the proper domain for kin selection theory. Ultimately, subjective factors determine how different people choose to split domains and attach labels. If someone chooses to associate genetic correlations with 'kin selection' and nongenetic associations with 'correlated selection', that is fine as long as the choice is expressed clearly and understood as a subjective choice.

In the absence of analysing particular problems, I would be inclined to separate kin selection and correlated selection. Those processes seem different, and so it makes sense to differentiate between them. However, I have repeatedly found that separating in that way is very unnatural when actually analysing particular problems (Frank, 1998). Neither the mathematics nor selection distinguishes the way in which phenotypic correlations between social partners arise. For example, in problems that follow the structure that led to eqn 10, the causal effect captured by the phenotypic correlation r depends only on the phenotypic association. The genetic aspects of transmission are handled independently by the τ terms.

For the phenotypic associations, one could choose to separate the causes of association into shared genotype and other factors. That separation would distinguishes between a narrow interpretation of kin selection and a residual component of correlated selection. That separation can certainly be useful. But repeatedly, in analysing particular problems and in developing the underlying abstract theory, the mathematics unambiguously leads one to blur the distinction when focusing on the causal analysis of how selection shapes phenotypes. For the particular causal component that concerns differential success separated from transmission, it is only the phenotypic correlations that matter.

Intuition often runs against the lessons urged by logical and mathematical analyses. That discord is perhaps the most interesting aspect of mathematics. People tend to split over that discord. Most trust their intuition above all else. Some, having felt the failure of their intuition too many times in the face of unambiguous logic, give in to the mathematics. I follow the latter course, in which it is much better to adjust intuition to mathematics and logic than to try and bend mathematics and logic to fit intuition.

In my view, the mathematics of selection has led inevitably to certain developments in the theory. Over time, the theory came to subsume the early ideas of kin selection into a broader causal perspective. That broader perspective is much more powerful when trying to analyse particular problems, and much simpler and conceptually deeper when trying to grasp the fundamental principles of evolutionary change. However, tastes vary. Others will prefer to separate and label differently. If one properly understands the underlying theory, different labelling causes few problems and ultimately is not a particularly interesting issue.

Direct and inclusive fitness

Consider two alternative ways to calculate fitness. The direct fitness method counts only the direct reproduction of individuals. If an individual behaves altruistically, we count only the negative effect of that behaviour on the individual. If that individual's social partners behave altruistically, then we add to the direct reproduction of our focal individual the benefit received from the altruism of neighbours. To calculate the total effect over the whole population, we sum up all of the positive and negative effects on the direct fitness of each individual, based on the individual's own phenotype, the phenotype of each individual's social

partners, and the heritabilities through different pathways of transmission.

Inclusive fitness alters the assignment of fitness components. If an individual behaves altruistically, we assign two components of fitness to that individual. The negative cost of altruism reduces the individual's own reproduction. The benefit of altruism to neighbours increases the neighbours' reproduction. We assign that increase in neighbours' fitness to the original altruistic individual that caused that increase, rather than directly to the neighbours themselves. We discount the neighbours' fitness component by their genetic similarity to the altruistic individual. Thus, the individual that expressed the behaviour is assigned both the direct effect on its own fitness and the indirect effect on neighbours' fitness discounted by genetic similarity.

Hamilton's approach

Hamilton's mathematical analysis showed that, under some conditions, inclusive fitness provides the same calculation as direct fitness. Hamilton preferred inclusive fitness, because it assigns all fitness changes to the behaviour that causes the changes. Some of the fitness changes are direct effects on the individual expressing the behaviour, and some of the fitness changes are indirect effects on other individuals receiving the behaviour. This assignment of all fitness effects back to the behaviour that caused them provides a clearer sense of cause and effect. Clear causal analysis aids in reasoning about the evolution of complex social behaviours. For example, inclusive fitness emphasizes that the effects of a behaviour on the reproduction of passive recipients can play a key role in determining whether genes associated with the behaviour tend to increase in frequency.

Hamilton understood that direct fitness was the ultimate measure for evolutionary analysis. His mathematical studies primarily had to do with showing that inclusive fitness was equivalent to direct fitness under many conditions. Hamilton emphasized inclusive fitness as his primary contribution to understanding social evolution. He discussed how inclusive fitness should be regarded as the fundamental process that encompasses kin selection, group selection, and other approaches to social interactions between genetically similar individuals (Hamilton, 1975). Almost all debates about the costs and benefits of Hamilton's approach and descendant ideas focus on inclusive fitness.

Development of the theory and failure of inclusive fitness

Since Hamilton's initial work, the study of social evolution expanded to analyse a broader and more realistic range of evolutionary problems. In my view, inclusive fitness has become as much a hindrance as an aid to understanding. I am not saying that inclusive fitness is wrong. Inclusive fitness does provide significant insight into a wide variety of problems. But one must know exactly its limitations, otherwise trouble is inevitable. Realistic biological scenarios arise for which inclusive fitness is important but not sufficient. When one does not clearly recognize the boundaries then, when faced with a solution for which inclusive fitness is not sufficient, it becomes too common to conclude that inclusive fitness and all broader approaches to kin selection analysis fail entirely, and one must discard the whole theory.

The issues are somewhat technical in nature. I provided a full analysis and discussion in Frank (1997a, 1998). Here, I give a sense of the problem and why it matters. I begin by briefly summarizing the main points from the previous section, which distinguished alternative measures of association between individuals. Separating those different kinds of association must be done clearly in order to understand the distinction between direct to fitness and inclusive fitness.

In the previous analysis leading to eqn 10, two different classes of individuals interacted. Consider first the direct fitness of class one individuals. They lose the cost C for their altruistic behaviour. Their social partners from the same class express an altruistic behaviour that provides an average benefit rB to a member of the class. The B is the beneficial trait of partners per unit of costly phenotype expressed by each individual, and r is the phenotypic association between the costly behaviour of an individual and the beneficial phenotype of partners. Thus, the total direct fitness effect on each individual of class one is proportional to $rB - C$. The heritability of the altruistic phenotype for class one individuals is τ_1, thus the heritable increase in altruism from the direct effect of class one individuals is $(rB - C)\tau_1$.

A second class of individuals does not express the altruistic phenotype, but may carry genes for that phenotype – for example, genetic relatives that receive care but do not give care. The net beneficial effect of altruism from class one on the direct fitness of class two individuals is $\hat{B}$. The heritability for class two individuals of the altruistic trait expressed by class one individuals is τ_2. Thus, the total heritable increase in altruism through the direct reproduction of class two individuals is $\hat{B}\tau_2$. Putting the direct fitnesses of the two classes together and weighting them equally leads to eqn 10 from the previous section, repeated here for convenience

$$(rB - C)\tau_1 + \hat{B}\tau_2 > 0.$$

Note the two different kinds of association, r and τ. The r coefficient measures the phenotypic association between the altruistic expression in social partners from class one. It does not matter how that phenotypic association arises. It may be caused by shared genotype, in

which case it is a common type of genetic relatedness coefficient. Or it may be caused by shared environment, such as sunlight or temperature, that is independent of genotype. No matter the cause of the phenotypic association, the direct fitness of class one individuals is proportional to $rB - C$. The actual value of r is a regression coefficient, and is sometimes called a coefficient of relatedness. However, it is more general than a coefficient of relatedness, because many different kinds of causes may be involved. With regard to immediate evolutionary consequences, the cause of the association does not matter.

By contrast, the τ coefficients measure heritability, and so can reasonably be understood as a measure of genotypic contribution to the expression of the altruistic character. In the case of τ_1, the measure is the heritability through the direct reproduction by an individual that expresses the altruistic behaviour. The τ_2 coefficient measures the direct contribution of class two individuals to the increase in the altruistic character, even though those individuals do not express the character. Because $\hat{B}$ represents an increment in fitness caused by the behaviour of class one individuals, the heritability of the altruistic phenotype expressed by class one individuals through the increment of fitness in class two individuals is proportional to the shared genotype between the class one actors and the class two recipients.

If class two recipients are genetically unrelated to class one actors, then $\tau_2 = 0$, and the condition for the increase in altruism is $rB - C > 0$. That has the form of Hamilton's rule. However, r measures phenotypic association, no matter the underlying cause. It may be that social partners in class one are genetically unrelated but phenotypically associated. Nonetheless, $rB - C > 0$ is still the proper condition, although it is certainly not an inclusive fitness expression in the manner usually understood by that theory. One can adjust definitions so that inclusive fitness still works. But the clearest understanding comes from analysing direct fitness, so that r carries its natural interpretation as a phenotypic association that may becaused by shared genes or may be caused by some other shared nongenetic process.

Alternatively, suppose that the phenotypic association between social partners in class one is zero, $r = 0$. Then, the condition for the increase in altruism is

$$-C\tau_1 + \hat{B}\tau_2 > 0.$$

Now consider the interpretation of the τ coefficients in terms of transmission and heritability. The ratio of indirect heritability to direct heritability is $R = \tau_2/\tau_1$. That coefficient, R, is the form of genetic relatedness commonly used in inclusive fitness theory. For inclusive fitness, one measures the relative transmission of causal genes through indirect compared with direct pathways of reproduction, which is the ratio of heritabilities. If, in the prior expression, we divide by τ_1, and use $R = \tau_2/\tau_1$, then we have

$$RB - C > 0,$$

in which R is the inclusive fitness coefficient of relatedness. This form is the classic expression of Hamilton's rule, which we may interpret with respect to inclusive fitness.

The direct fitness approach gives the correct analysis in all cases, with proper interpretation of r as a phenotypic association between social partners and τ as transmission to the future through heritability. Inclusive fitness arises as a special case. By contrast, if one begins with an inclusive fitness perspective, one has to struggle to obtain the right interpretation, and confusion will often arise with regard to both the analysis and the interpretation.

The actual distinctions between direct and inclusive fitness are more extensive and more subtle (Frank, 1998, Chapter 4). Direct fitness typically provides a clear and complete analysis, and subsumes inclusive fitness as a special case. Inclusive fitness does have the benefit of an intuitively appealing causal perspective. However, inclusive fitness is more limited and more likely to cause confusion. As understanding of a subject develops, it is natural for yesterday's general understanding to become today's special case.

Understanding how selection shapes phenotypes

Hamilton (1970) originally set out to develop a causal decomposition of social evolution into components. His decomposition by inclusive fitness had two steps: the separation of fitness into components and the analysis of heritability. With regard to fitness, Hamilton's approach partitioned the total effect of a phenotype into the direct consequence on the actor and the indirect consequence on social partners. With regard to heritability, Hamilton weighted the different fitness components by their fidelity of transmission relative to the phenotype in the focal individual. His coefficient of relatedness measured the ratio of the heritability through the indirect fitness component of social partners relative to the heritability through the direct reproduction by the actor.

Multivariate selection and heritability

Independently of Hamilton's work, the theory of natural selection developed during the 1980s and 1990s. That development primarily followed the influential paper by Lande & Arnold (1983), which built on two earlier lines of thought. First, Pearson (1903) had established the partitioning of fitness into distinct components. Second, Fisher (1918) had established the modern principles of heritability and the conceptual foundations of quantitative genetics.

The development of these two lines – the components of fitness and the components of heritability – independently paralleled the two lines in Hamilton's thought. In retrospect, the parallel development is not surprising. Anyone attempting a causal analysis of selection and evolutionary change would ultimately be led to those same two essential parts of the problem.

In the early 1980s, I began to develop my own methods for analysing phenotypes influenced by kin selection. I started with the Price equation methods that I inherited from Hamilton's graduate seminar in 1979. As I refined my methods of analysis and then eventually generalized the approach with the help of Peter Taylor, I was inevitably up against the two problems of partitioning fitness into components and tracing pathways of heritability. However, until my work with Taylor in 1996, I had not given much thought to the underlying structure of the problem.

Following 1996, when I found Queller's papers that merged Hamilton's kin selection theory with the Lande and Arnold method of multivariate selection and quantitative genetics, I began work on developing that connection identified by Queller. The result is that kin selection and inclusive fitness became part of the broader approach to the study of natural selection (Frank, 1997b, 1998, 2012b,c, 2013). With that advance, it is no longer possible to separate cleanly between the initial view of kin selection as a special kind of social problem among genetically similar individuals and the broader approach of causal analysis for phenotypes, fitness and heritability.

From the merging of kin selection theory and the broader aspects of selection and heritability, problems like the analysis of altruism between species have come to look like a kin selection analysis, and classical problems of kin selection have come to look like a Lande & Arnold type of analysis of multivariate selection, with the addition of a more complex analysis of heritability.

Statics and the three measures of value

With regard to studying particular biological problems, I continue to favour comparative statics for its pragmatic approach. In analyses of comparative statics, kin selection problems are transformed into the analysis of three measures of value: marginal value, reproductive value and the valuations of relative transmission (Frank, 1998).

Marginal values transform different phenotypic components into common units. Suppose, for example, that we analyse the marginal costs of a behaviour associated with the direct reproduction of an actor and the marginal benefits of that behaviour associated with the indirect effect on the fitness of a social partner. The relative marginal valuations provide a substitution, or translation, measure. That measure tells us, for each small change in phenotype, how much the marginal benefits change relative to how much the marginal costs change. For example, does a small change in the costs for direct reproduction translate into a small or a large change in the benefits for social partners? It is the relative marginal valuations that give us that translation.

Marginal valuation only applies to the analysis of small changes. More generally, when one analyses large changes, the regression coefficients from multivariate analysis arise. In that context, the regression coefficients serve as translations for the relative scaling between different phenotypes and components of fitness (Frank, 2013).

Reproductive value provides a weighting for different kinds of individuals with respect to their contribution to the future of the population. Reproductive value is a component of the transmission of phenotypes. However, we separate reproductive value from heritability, because reproductive value usually differs by demographic rather than genetic aspects. For example, age is a demographic property, and individuals of different ages have different reproductive values, although they may have the same heritability in transmission of their phenotypes. Ecological factors, such as available resources or the tendency for local extinctions of groups, also influence reproductive valuation. In terms of classical demography, resource availability may affect birth rates and local extinctions of groups may affect death rates.

Valuations of relative transmission, or heritability, obviously play an essential role in tracing the causes of evolutionary change by natural selection. The coefficients of relatedness in the initial theories of kin selection had to do with relative heritabilities through different pathways of transmission. Those coefficients of relatedness are just special cases of the broader analysis of heritabilities in the general study of natural selection.

Social evolution and traditional kin selection problems raise particular issues with regard to the three measures of value. However, the analysis of those social aspects falls within the broader framework of natural selection, which applies to all problems of selection and the transmission of phenotypes to future generations. This merging of kin selection and inclusive fitness into the broader framework for the study of selection has led to deeper understanding and more powerful analytical approaches. At the same time, the separate initial history of kin selection compared with the analysis of nonsocial problems sometimes leads to confusion about the current understanding of the theory of social evolution.

The failure of group selection

> Properly understood, then, the origins of an idea can help to show what its real content is; what the degree

> of understanding was before the idea came along and how unity and clarity have been attained. But to attain such understanding, we must trace the actual course of discovery, not some course which we feel discovery should or could have taken, and we must see problems (if we can) as the men of the past saw them, not as we see them today.
>
> In looking for the origin of communication theory one is apt to fall into an almost trackless morass. I would gladly avoid this entirely but cannot, for others continually urge their readers to enter it. I only hope that they will emerge unharmed with the help of the following grudgingly given guidance (Pierce, 1980, pp. 20–21).

A causal analysis of selection begins by expressing how phenotypes and other variables influence the fitness of individuals. In social problems, the characteristics of an individual's local group sometimes enter the expression for individual fitness. If group characters influence fitness, then a causal component of selection is attributed to the group. That causal component attributed to the group is one common way in which group selection arises.

Hamilton developed models of group selection by this partition into individual and group characters. I used Hamilton's group selection methods in my own early studies. Later, I came to understand the limitations and ultimate failure of the group selection perspective. I then merged kin selection theory with the general causal analysis of selection and transmission.

Hamilton's group selection models

In the 1970s, Hamilton studied social phenotypes in group structured populations. He analysed sex ratios and dispersal polymorphisms of wasps that live in figs (Hamilton, 1979). Each fig formed a clearly defined group. He also studied multigeneration groups of insects and other arthropods that lived in isolated rotting logs (Hamilton, 1978). As always, Hamilton combined his natural history observations with mathematical models to analyse natural selection. For these group structured problems, he followed the hierarchical multilevel methods of Price (1972), as described in Hamilton (1975).

Interestingly, a Price equation analysis of group structured populations is similar to a Lande & Arnold analysis of multivariate selection. In the case of group structure, fitness depends on an individual's phenotype and on the average phenotype of social partners in the group. That decomposition of fitness into individual and group components, when used in the Price equation, gives a causal decomposition that ascribes effects to the individual and group phenotypes. The causal component attributed to groups may be interpreted as group selection. There is nothing special about using individual and group phenotypes in a Price equation analysis of fitness. If one used an individual's phenotype, the phenotype of the individual's mother, and temperature, one would obtain a decomposition in terms of those variables. The analysis works for any choice of variables that affect fitness.

Lande & Arnold (1983) also used the Price equation for their analysis of multivariate selection. Lande & Arnold's approach was in fact very similar to the unpublished Price equation method Hamilton used to analyse the sex ratios of fig wasps. However, Hamilton did not interpret his Price equation method broadly as the multivariate analysis of selection, but instead followed Price's limited interpretation of partitioning individual and group components of success.

Following Hamilton, I began my own studies of dispersal and sex ratios by thinking in terms of group structured natural history. The mathematical models partitioned individual and group components of fitness (Price, 1972; Hamilton, 1975). Hamilton was not a committed group selectionist in the sense that began to develop in the 1980s. Instead, Hamilton interpreted group structure as one way to obtain a positive genetic association between individuals, as emphasized very clearly in the quotes from Hamilton (1975) that I presented in an earlier section. To some extent, Hamilton's strong focus on group structure arose from his inability to analyze phenotypes such as dispersal and sex ratios in terms of kin selection and inclusive fitness. He understood that those processes were the key, but he could not write down mathematical analyses in terms of kin selection. He had access to Price's methods for group structuring and so used that method instead.

Hamilton was not fully satisfied with his group level analysis of sex ratios as given in his 1979 notes from his graduate course at the University of Michigan. He never published that analysis, perhaps because it showed only that greater genetic similarity within groups led to a stronger kin selection effect. That vague point was already obvious, as he had emphasized in his 1975 article. Simply to show that vague conclusion again for sex ratios did not add any real insight.

My early group selection models

I took up the empirical study of fig wasp sex ratios in 1981. At that time, I also began to study Hamilton's notes and to learn how to extend Price's hierarchical multilevel selection analysis to apply to my empirical work. My initial success with that method was limited to seeing that greater group structuring and more limited migration increased the genetic similarity of individuals within groups. The more closely related group members are, the stronger the kin selection effects.

By the reasoning from Hamilton's, 1975 paper and his teaching, I understood the equivalence between the kin selection perspective and the conclusion that the greater genetic variance among groups, the stronger the tendency for biased sex ratios. So, one could call that a group selection perspective. But there was always the clear understanding that the ultimate causal basis arose from kin selection or inclusive fitness perspectives. I summarized my early understanding of hierarchical multilevel selection and related group selection analyses in my first article, *A hierarchical view of sex ratio patterns* (Frank, 1983a).

In my early work, I focused only on group structured populations. I maintained an unbiased dual perspective between kin and group selection. However, in that early work, I made limited progress towards teasing apart how selection influenced sex ratio evolution. I was stuck at the same vague group selection perspective that stopped Hamilton. I slowly figured out how to move ahead. Particular sex ratio models played a key role – the synergism between application and abstraction.

In a particular study, I analysed the case in which males competed for mates locally within groups, females competed for resources against neighbouring females, and the males and females migrated varying distances before mating. I then traced the causal processes that determined the evolution of the sex ratio. Assuming that the mother controlled the sex ratio of her progeny, one could adopt the mother's perspective with regard to pathways of causation. This new work grew from Price and Hamilton's multilevel selection analyses, reflected in the title of a key article, *Hierarchical selection theory and sex ratios. I. General solutions for structured populations* (Frank, 1986c). Although that article emphasized hierarchical multilevel selection, it also placed the group structured perspective into its proper role: a special case within the broader analysis of phenotypes by kin selection theory.

Pathways of causation replace group selection

In the group structured sex ratio models, one can separate several distinct causes with respect to kin interactions. For example, the value of an additional son depends on the mother's genetic relatedness to the males that her sons compete with for mates. Greater relatedness reduces the transmission benefit to a mother for an additional son. We can express the effect as the marginal gain in mating success through an additional son multiplied by the relative heritability of the mother's sex ratio trait through sons minus the marginal loss in mating success among competing males multiplied by the heritability of the mother's sex ratio trait through those competing males.

The value of an additional daughter depends on the mother's genetic relatedness to the females that her daughters compete with for access to resources. Greater relatedness reduces the transmission benefit to a mother for an additional daughter. We can express the effect as the marginal gain in reproductive success for an additional daughter multiplied by the relative heritability of the mother's sex ratio trait through daughters minus the marginal loss in reproductive success among competing females multiplied by the heritability of the mother's sex ratio trait through those competing females. In addition to the direct contributions through each sex, there are also effects of one sex on the other. For example, an extra daughter may provide additional mating opportunities for sons.

The full analysis showed how various causal pathways influence the predicted sex ratio (Frank, 1985b, 1986b,c). Those pathways often include the genetic associations between competitors, measured by coefficients of relatedness. Typically, a coefficient of relatedness can be expressed either as the genetic variance between groups divided by the total genetic variance in the population, or as a regression coefficient that measures the genetic correlation between interacting individuals. The two interpretations are simply alternative expressions for the same measure. The first, group based expression for the measure suggests a group selection interpretation, whereas the second, individual-based expression suggests a kin-centric interpretation. However, the measure is the same in both cases (Frank, 1986c, 1998).

The problem with the group-based interpretation is that different causal pathways may be associated with different patterns of grouping. Or there may not be any natural grouping. The pairwise correlations of kin selection theory do not require group structure. If there is no group structure, kin selection works perfectly whereas group selection fails.

Group selection, which initially provided a nice intuitive way to think about group structured populations, ultimately proved to be the limitation in understanding the evolution of phenotypes. Inevitably, I had to return to the fundamental causal level, in which the correlations between individual phenotypes and the different pathways of heritability were made clear.

Once at the proper level of causation, one could see that emphasis on groups hindered analysis. The proper view always derives from the causes of fitness and the pathways of transmission. The different causes of fitness rarely follow along a single pattern of grouping. For example, males may interact over one spatial scale, females may interact over another spatial scale, and mating between males and females may occur on a third scale. It is relatively easy to trace cause by the correlations between phenotypes, the ecological context of resource distribution and the pathways of genetic transmission. Those causal components do not naturally follow the sort of rigid grouping needed for a group selection analysis to work.

The last paragraph of Frank (1986c) emphasized how analysis of particular biological problems led to a deeper understanding of causal process:

> In summary, all three theories – inbreeding, within-sex competition among relatives, and group selection truly describe causal mechanisms of biased sex ratios in structured populations. Through the study of a variety of scenarios with hierarchical selection theory, I draw the following conclusions. First, inbreeding biases the sex ratio since producing a daughter that inbreeds … passes on twice as many parental genes as producing a son would. Second, as the amount of within-sex competition among related individuals increases, the relative genetic valuation of that sex decreases. Third, genetic differentiation among groups … and genetic correlation within groups … are related descriptions for the same phenomenon. Some recent papers (Colwell, 1981; Wilson & Colwell, 1981) have stressed the group selection aspect of this phenomenon without clarifying its similarity to genetic relatedness. Using group selection for describing causal mechanisms is particularly slippery, since, as in the various scenarios presented in this paper the differentiation among groups may refer to groups of competing males, groups of competing females or groups that contain inbreeding pairs. While hierarchical selection theory, which is a group selection sort of analysis, has proved a powerful analytical tool, it seems that, for describing causal mechanisms, it is often useful to apply the genetic regressions [kin selection coefficients] considered in the discussion.

Logically, there cannot be a group selection controversy

Two conclusions emphasize the failure of group selection. First, the ultimate causal processes concern correlations between phenotypes and pathways of genetic transmission. Group structuring is just one limited way in which phenotypic correlations and genetic transmission pathways may be influenced. Second, insisting on a group perspective greatly limits the practical application of the theory to natural history. Most natural history problems do not have a single rigid group structure shared by all causal processes. If one starts with a group selection perspective, solving problems becomes extremely difficult or impossible. No gain in understanding offsets the loss in analysis.

Although group selection has problems that limit its scope, it also has attractive features. There is a natural intuitive simplicity in group structured analysis. Total selection arises from the balance between the dynamics of selection within groups and the dynamics of selection between groups. Altruistic characters often tend to lose out during selection within groups and often tend to increase by selection between groups. The problem is that once people gain such intuition, they do not easily give it up in the face of the inevitable conceptual and practical limitations. Like the growth of any kind of understanding, one must allow the first general insights to become the special cases of broader conceptual and analytical approaches. Pinning a topic to the first simple illustrative model limits progress. Concepts and their associated language naturally develop and transform over time.

Given this history, the idea of a group selection controversy seems to me to be a logical absurdity. I do understand the intuitive appeal of group selection. I was trained by Hamilton to think about the interesting properties of groups and about the dynamical processes of within-group and between-group selection. I also learned from Hamilton the mathematical techniques to analyse multilevel selection. My early conversion, however, did not last. Both the conceptual and practical limitations became apparent as I tried to make progress in understanding various problems of natural history and the mathematical models needed to evaluate those problems.

In summary, when groups are the cause of genetic associations, then group structuring is the causal basis for the associations that drive kin selection. When group structuring is less clear, the principles of kin selection still hold, as they must.

However, the controversy continues

The Los Angeles Times newspaper published an interview with E.O. Wilson on 19 September 2012. With regard to kin and group selection, the interviewer began:

> The biologist J.B.S. Haldane explained 'kin selection' when he was asked whether he would lay down his life for his brother. No, he said, but he would for two brothers, or eight cousins. In the journal 'Nature' in 2010 (Nowak *et al.*, 2010), you challenged kin selection and created a stir, to say the least.

Wilson replied (the following is an exact unmodified typographic transcription of the published article)

> I was one of the main promoters of kin selection back when it looked good. By the '90s I thought I heard the whine of wheels spinning. Willie Hamilton's [*British evolutionary biologist W.D. Hamilton*] generalized rule [of kin selection] was that if you've got enough people looking after [relatives], society could become very advanced. It wasn't working. By 2010, I had published peer-reviewed articles on what was thoroughly wrong [with kin selection]. I said we've got to go back to 'multilevel selection'. Groups form, competing with one another for their share. It's paramount in human behaviour. The spoils tend to go to groups that do things better – in business, development, war and so on.

> I knew the biology. I saw that multiple-level selection works, but in different ways in different cases, and [with] my mathematical colleagues, said [in the Nature article], kin selection cannot work. We knew that was going to be a paradigm changer. We published it and the storm broke.

I agree with the three key points emphasized by Wilson in his article (Nowak *et al.*, 2010): that multilevel selection is important; that one should think about group selection in human sociality; and that kin selection in relation to haplodiploidy is not sufficient to explain insect sociality. However, I do have a different perspective on some of the conceptual issues and the history of the subject.

Multilevel selection

I have emphasized Hamilton's own interest in multilevel selection. Thus, Wilson's way of opposing kin selection vs. multilevel selection does not make any sense to me. To expand briefly on this point with regard to human sociality, note that Hamilton's (1975) primary publication on multilevel selection had the title *Innate social aptitudes of man: an approach from evolutionary genetics*. In that article, Hamilton first developed his theoretical perspective on human evolution by extending Price's (1972) hierarchical selection methods. Hamilton then devoted approximately ten pages to group level perspectives on human sociality.

In Hamilton's (1996) collected works, he gave the reprinted version of this article (Hamilton, 1975) the secondary title *Friends, Romans, Groups*, and wrote in the preface for the article:

> [I] am proud to have included the first presentation of Price's natural selection formalism as applied to group-level processes. … He [Price] himself published one application of the formula to groups but I think it was less explicit and general than mine, indeed almost as if he was trying still to conceal his formula's full significance (Price, 1972). For myself, I consider the format of analysis [for multilevel selection] I was able to achieve through his idea brilliantly illuminating.

Kin vs. group selection in human sociality

Alexander (1979, 1987) built his comprehensive evolutionary analysis of human sociality on the importance of group against group competition. In developing the theoretical foundations for his analysis, Alexander had thoroughly reviewed issues of group selection. He expressed his thinking on this topic in an article with the title *Group selection, altruism and the levels of organization of life* (Alexander & Borgia, 1978).

I took my first undergraduate course in evolution from Alexander at the University of Michigan in 1978. His views on multilevel selection in Alexander & Borgia (1978) were particularly important in shaping how I understood the subject. Interestingly, Alexander was not much influenced by Hamilton's multilevel selection analysis, which derived from the mathematical theories of Price. Instead, Alexander was an entomologist with a deep interest in human behaviour. He developed his thinking from broad consideration of natural history.

A comment on the first page of Alexander & Borgia (1978) provides historical context

> [E]volution by differential extinction of groups has recently been modelled or discussed anew by several authors … Wilson (1975b), for example, has argued that 'In the past several years a real theory of interpopulation selection has begun to be forged, with both enriched premises and rigorous model building. … Insofar as the new theory considers the results of counteraction between group and individual selection, it will produce complex, nonobvious results that constitute testable alternatives to the hypothesis of individual selection. My own intuitive feeling is that interpopulation selection is important in special cases'.

Clearly, Wilson has long given thought to the potential importance of group structuring in nature. The point here is that this line of thought goes back over 40 years, with much debate about conceptual issues and the relative importance for understanding natural history. However, most theories conclude that differential extinction of groups is usually a relatively weak force (Maynard Smith, 1976). Only Hamilton's milder version of group structuring in relation to differential reproduction and genetic differentiation between groups seems to be on solid ground as an explanation for common patterns of natural history.

I have argued throughout that Hamilton's type of multilevel selection is clearly a special case within the broader theory of kin selection. Although group extinctions are usually thought to be outside the scope of kin selection theory, the merging of demography and kin selection by Taylor & Frank (1996) and Frank (1998) brings those different processes within a single coherent theoretical framework. Bringing together those different processes is more than just a theoretical convenience.

Suppose, for example, that a particular problem requires analysing how an altruistic phenotype evolves. That phenotype may affect the probability of extinction for its group. Its group may be composed of genetically similar individuals, perhaps kin in the traditional sense. The phenotype may also have other costs and benefits with regard to interacting with social partners. It is too hard to figure out what to expect by separately analysing extinctions, group variations in genetics, social processes of costs and benefits, and other processes. Nowak *et al.* (2010) say that one must make a specific model for a specific case, and then one gets the right answer. True, but the history of science shows unambiguously that one gains a lot by understanding the abstract

causal principles that join different cases and different models within a common framework (Frank, 2012b). That truth leaves only the sort of causal perspective I have emphasized as a reasonable candidate for how to think about such problems.

Certainly, not everyone agrees with me. Alexander & Borgia (1978) understood Hamilton's (1975) claim that group selection was just a special case of kin selection. However, they rejected that point of view

> That groups are often composed of kin does not mean that kin selection and group selection are in any sense synonymous (e.g. Williams, 1966; West-Eberhard, 1975; Wilson, 1975a,c). As West-Eberhard (1976) points out, 'In the same trivial sense that kin selection is group selection, all of natural selection is group selection, since even "individual" selection really concerns the summed genetic contribution of a group – the individual's offspring'. Moreover, although kin selection can occur in continuously distributed populations, group selection cannot. For reasons elaborated later, we agree with Maynard Smith (1971) that it is more appropriate to distinguish kin selection and group selection than to blur their differences by considering them together.

This quote shows the clear historical precedent for Wilson's argument that multilevel selection is distinct from kin selection. As often happens with historical analyses of ideas in science, one can find significant antecedents that support a variety of different positions. Because a controversy about kin and group selection will always come down to how one chooses to interpret words, there can be no final resolution.

What we can say, unambiguously, is that Hamilton never argued for a distinction between multilevel analysis and kin selection. Instead, he saw multilevel analysis as one of the most powerful approaches to thinking about the general problems that arise in applications of his theory. My own view follows Hamilton. In addition, the mathematics do not allow a logical distinction. Any distinction must be injected by a particular bias with respect to how one uses the words and interprets the history. However, as long as the historical and conceptual issues are made clear, it does not matter to me how one chooses to use the words. Indeed, given how clearly we understand the theory, it puzzles me why so much attention and argument continues to focus on this issue.

Returning to Alexander & Borgia (1978), their main point concerned how we should think about human evolution. In their concluding remarks about humans, they state

> Human social groups represent an almost ideal model for potent selection at the group level (Alexander, 1971, 1974, 1975; Wilson, 1973). First, the human species is composed of competing and essentially hostile groups that have not only behaved towards one another in the manner of different species but have been able quickly to develop enormous differences in reproductive and competitive ability because of cultural innovation and its cumulative effects. Second, human groups are uniquely able to plan and act as units, to look ahead, and to carry out purposely actions designed to sustain the group and improve its competitive position, whether through restricting disruptive behaviour from within the group or through direct collective action against competing groups.

Alexander & Borgia (1978) certainly understood the broader implications of multilevel selection analysis with regard to a variety of biological problems. The first paragraph of their summary is

> [T]here may be few problems in biology more basic or vital than understanding the background and the potency of selection at different levels in the hierarchies of organization of living matter. The approaches currently being used by evolutionary ecologists and behaviorists in assessing the likelihood of effective selection at the level of groups or populations of individuals may also be used to advantage by those concerned with function at intragenomic levels. The kind of selectionist techniques used recently to analyze the behavior of nonhuman organisms may in the near future be widely applied toward understanding not only human social phenomena, but a variety of phenomena of classical biology such as mitosis, meiosis, sex determination, segregation distortion, linkage, cancer, immune reactions, and essentially all problems in gene function and in ontogeny.

The year 1978 was a long time ago. I do not understand why these ideas are still thought of as novel or controversial.

Insect sociality. Wilson particularly emphasized the failure of kin selection theory in explaining the evolution of advanced sociality in insects (Nowak *et al.*, 2010). Once again, we must consider what is meant by the scope of kin selection theory. I tend to think of kin selection as a particular causal perspective within the broader theory of natural selection. Although it is tempting to limit the scope of kin selection theory to certain simple scenarios and predictions, such limits never make sense in the logical or mathematical analysis of the subject. However, to the extent that others choose to distinguish more finely, that does not bother me as long as the concepts and history are made clear.

For those who view kin selection narrowly, the potential problems of that narrow theoretical view for understanding the origin of complex sociality (eusociality) in insects was a lively topic in the 1970s and 1980s. Andersson (1984) introduces his excellent review of the topic by noting that eusociality has arisen multiple times in insects. He then turns to an early the-

ory based on kin selection to explain why eusociality is particularly common in bees, ants and wasps (Hymenoptera).

> Hamilton's (1964a,b) celebrated explanation is that haplodiploid sex determination in Hymenoptera makes sisters share three quarters of their genes, whereas a daughter only receives half her genome from her mother. Hymenopteran females may therefore propagate their genes better by helping to raise reproductive sisters than by raising daughters of their own. Haplodiploidy therefore should make the evolution of nonreproductive female workers particularly likely among the Hymenoptera. This and other stimulating ideas of Hamilton's started a revolution in the study of social behaviour, particularly of the role of kin selection (Maynard Smith, 1964; Michod, 1982).
>
> Several entomologists have warned against overemphasis on the 3/4 relatedness hypothesis, and they have pointed to other factors important in the evolution of eusociality (e.g. Kennedy, 1966; Lin & Michener, 1972; Alexander, 1974; Michener, 1974; Michener & Brothers, 1974; West-Eberhard, 1975, 1978; Evans, 1977; Crozier, 1979; Eickwort, 1981; Brockmann, 1984). Hamilton (1964a,b, 1972) and Wilson (1971) also noted that haplodiploidy alone cannot explain eusociality in Hymenoptera. Such reservations were often forgotten, however, and the 3/4 hypothesis came to dominate many textbook and popular accounts. For example, in his comprehensive review of social behaviour in animals, Wilson (1975c, p. 415) stated that 'the key to Hymenopteran success is haplodiploidy' and that 'nothing but kin selection seems to explain the statistical dominance of eusociality by the Hymenoptera' (Wilson, 1975c, p. 418). A long list of similar evaluations of the 3/4 relatedness hypothesis by other authors could be cited.
>
> The main empirical evidence in favour of the 3/4 hypothesis is that eusociality seems to have arisen many more times in the haplodiploid Hymenoptera than in other insects (Wilson, 1971; Brockmann, 1984). This evidence initially appeared impressive, but several recent findings indicate that haplodiploidy and 3/4 relatedness between sisters may have been of limited importance for the evolution of eusociality. Other factors have clearly been involved, and it seems possible that haplodiploidy has even been insignificant compared to these factors. At least five lines of evidence cast doubt on the overwhelming importance sometimes ascribed to haplodiploidy [and the narrowly defined kin selection hypothesis].

The further details do not concern us here. The main point is that by 1984, the problems with a narrow interpretation of kin selection for explaining eusociality had been widely discussed.

Summary. Inevitably, the debates about kin selection and group selection will continue, because the ultimate problem concerns different usage of words. People vary in whether they prefer to emphasize differences or similarities between components of a broader problem. Those who like differences emphasize distinctions between kinship interactions and group structuring of populations. Those who like similarities see kin and group selection as part of a broader theory of natural selection. So what? Perhaps, the debate can advance a bit by a more nuanced consideration of the underlying concepts and history.

Discussion

Separation into component causes

Kin selection theory analyses the evolutionary causes of social phenotypes. Causal analysis is not an alternative to other analyses, such as population genetics. Rather, causal analysis brings out the factors that one must emphasize to understand pattern. Why do phenotypes vary in the way that they do? What matters most? What factors should one focus on to make testable predictions?

Hamilton separated evolutionary change into three causes. A direct component affects the individual that expresses a phenotype – the cost. An indirect component affects social partners influenced by the focal individual's phenotype – the benefit. To combine those two components into the total evolutionary effect on the phenotype, one must adjust the indirect component to have the same units as the direct component. The adjustment translates the indirect component of change into an equivalent amount of direct change. That translation is often a genetic measure of similarity between the individuals affected directly and indirectly – the coefficient of relatedness.

Note the structure of causal analysis. The total change is what matters. To understand that total change, we separate it into parts that help to reason about the problem. Once we have a set of distinct parts, we have to combine those parts back into a common measure for total change. To combine properly, each part is weighted by a factor that translates into a consequence for total change. Separation into distinct causal processes leads to testable predictions about which causal components explain patterns of variation. Separation also highlights the common causal basis that unites previously unconnected problems within a common conceptual framework.

Causal analysis by kin selection is never an alternative to other analyses of total change. Rather, it is a powerful complement to other approaches. In practice, kin selection can be so powerful in analysis and so helpful in the conceptual framing of problems, that one does not need other complementary methods. However, determining the best methods always

depends on the particular goals. For example, in the study of alternative genetic assumptions and complex aspects of dynamics, population genetic models provide superior methods.

Hamilton's rule

Confusion over Hamilton's rule arises when it is not properly understood as a partitioning of causes. The rule is the partition of total change for a social phenotype into direct and indirect components. It does not make sense to consider whether the rule is true or false. Rather, following Hamilton, one thinks of the rule in two ways. First, is there a simple form for the partition of causes that matches the ultimate measure of total change, at least approximately and under particular conditions? If so, what are the proper definitions for the components? Second, how should we expand Hamilton's original causal partition for more complex problems?

Roughly speaking, Hamilton's original expression in terms of costs, benefits and genetic relatedness provided a useful partition that works for simple problems. However, as the theory was applied to more realistic problems, the associated causal analysis had to be extended. The modern theory of kin selection provides a more comprehensive causal analysis. Multiple direct and indirect components of fitness may occur. Costs and benefits are understood to depend on context. Relatedness coefficients and their generalization by multiple regression coefficients translate all fitness components into common units of total change. Separation between selection and transmission clarifies the distinctions between causal components.

Methods of analysis for solving problems have been developed to complement the causal decomposition. The limitations of the analytical methods and the causal decompositions are reasonably well understood. Causal decomposition and simplified analysis provide tools to enhance understanding rather than alternatives to more complex and detailed mathematical analyses of particular problems.

Limitations of inclusive fitness

Hamilton introduced inclusive fitness as a particular type of causal partition. Inclusive fitness assigns an indirect fitness effect through a social partner back to the behaviour that caused the fitness effect. For example, if an individual saves its sibling's life, that fitness benefit is attached to the individual who saved the life rather than the individual whose life was saved. That fitness benefit is discounted by the genetic relatedness of the savior to the sibling.

Inclusive fitness has the advantage of assigning changes in components of fitness to the phenotype that caused those changes. That causal decomposition can provide much insight into evolutionary process. The problem arises because that very particular form of causal partitioning is often equated with the entire theory of kin selection. Instead, it is much better to view kin selection as a general approach to the causal analysis of social processes. Inclusive fitness is a particular causal decomposition that helps in some cases and not in others.

For example, phenotypic associations between social partners that do not share a common genotype can have a very powerful effect on social evolution. Inclusive fitness fails as a complete analysis of correlated phenotypes between social partners. That failure does not mean we should give up on trying to understanding the causes of social evolution in such cases, or that we should conclude that kin selection theory fails as a general approach. Instead, we must understand the broader approach of causal analysis, and how different aspects of natural history should be understood from a broader causal perspective. That broader perspective was developed many years ago and has proved to be a powerful tool for analysing complex social interactions.

Correlated social phenotypes vs. genetic relatedness

When correlated social phenotypes do not arise from shared genotype, how should we think of the relatedness coefficients of kin selection theory? Proper causal analysis solves the problem. Changes in phenotypes cause changes in fitness. Those fitness changes must be translated into changes in the transmission of phenotypes to the future population. In analysing the causes of fitness and the causes of transmission, we must put all the components together into a common measure of total change. The weighting of the different components leads to different types of regression coefficients. Those regression coefficients are the translations of different causal components into a common scale.

The fact that Hamilton's original theory considered only a very particular aspect of transmission and genetic relatedness has led to confusion. Hamilton's original regression coefficient of relatedness is not the single defining relatedness and regression coefficient of kin selection theory. Rather, it is the particular coefficient that arises in the special inclusive fitness analysis that Hamilton considered in developing his theory.

Synergism between abstraction and application

Abstraction arises by recognizing the common processes that recur in different cases. Application demands analysis of particular phenotypes under particular circumstances. Kin selection theory grew naturally by the synergism between abstraction and application. Hamilton pulled out the first clear abstraction that united various simple applications. Yet, he could not use his

abstract theory to move on to new applications. In particular, he could not solve the problems of dispersal and sex ratios that arose from kin interactions.

As the applied theory eventually developed for dispersal, sex ratios and more complex social phenotypes, deeper abstract principles emerged. For example, the distinction between selection and transmission became clear, and relatedness coefficients became a part of translating causal components into common units. The improvements in abstract theory enhanced the scope of application to complex social phenotypes.

The necessary synergism between abstraction and application showed the ultimate failure of group selection. In particular, group selection is a useful abstraction for a limited set of applications. When faced with a variety of applications, such as sex ratio evolution with multiple male and female interactions, group selection fails. Instead of the limited perspective of group selection, the deeper abstract principles dominate. Those principles include a clear causal analysis of distinct fitness components, separation of selection and transmission, and the proper weighting of the distinct causal components to attain an overall analysis of total change.

When different people focus exclusively on either abstraction or application, deep tension and fruitless debate arise. When the two modes come together, great progress follows.

Acknowledgments

National Science Foundation grant EF-0822399 supports my research.

References

Alexander, R.D. 1971. The search for an evolutionary philosophy of man. *Proc. R. Soc. Victoria* **84**: 99–120.

Alexander, R.D. 1974. The evolution of social behavior. *Annu. Rev. Ecol. Syst.* **5**: 325–383.

Alexander, R.D. 1975. The search for a general theory of behavior. *Behav. Sci.* **20**: 77–100.

Alexander, R.D. 1979. *Darwinism and Human Affairs*. University of Washington Press, Seattle.

Alexander, R.D. 1987. *The Biology of Moral Systems*. Aldine de Gruyter, New York.

Alexander, R.D. & Borgia, G. 1978. Group selection, altruism, and the levels of organization of life. *Annu. Rev. Ecol. Syst.* **9**: 449–474.

Alexander, R.D., Noonan, K.M. & Crespi, B.J. 1991. The evolution of eusociality. In: *The Biology of the Naked Mole Rat* (P.W. Sherman, J.U.M. Jarvis, R.D. Alexander, eds), pp. 3–44. Princeton University Press, Princeton, NJ.

Andersson, M. 1984. The evolution of eusociality. *Annu. Rev. Ecol. Syst.* **15**: 165–189.

Bijma, P. & Wade, M.J. 2008. The joint effects of kin, multilevel selection and indirect genetic effects on response to genetic selection. *J. Evol. Biol.* **21**: 1175–1188.

Bourke, A.F.G. 2011. *Principles of Social Evolution*. Oxford University Press, Oxford, UK.

Bourke, A.F.G. & Franks, N.R. 1995. *Social Evolution in Ants*. Princeton University Press, Princeton, NJ.

Brockmann, H.J. 1984. The evolution of social behaviour in insects. In: *Behavioural Ecology: An Evolutionary Approach*, 2nd edn (J.R. Krebs & N.B. Davies, eds), pp. 340–361. Blackwell Publishing, Oxford.

Burt, A. & Trivers, R. 2008. *Genes in Conflict: The Biology of Selfish Genetic Elements*. Belknap Press, Cambridge, MA.

Charlesworth, B. 1994. *Evolution in Age–Structured Populations*, 2nd edn. Cambridge University Press, Cambridge.

Cheverud, J.M. 1984. Evolution by kin selection: a quantitative genetic model illustrated by maternal performance in mice. *Evolution* **38**: 766–777.

Clark, A.B. 1978. Sex ratio and local resource competition in a prosimian primate. *Science* **201**: 163–165.

Colwell, R.K. 1981. Group selection is implicated in the evolution of female-biased sex ratios. *Nature* **290**: 401–404.

Crespi, B.J. 2001. The evolution of social behavior in microorganisms. *Trends Ecol. Evol.* **16**: 178–183.

Crozier, R.H. 1979. Genetics of sociality. In: *Social Insects*, Vol. 1 (H.R. Hermann, ed.), pp. 223–286. Academic Press, New York.

Darwin, C. 1859. *On the Origin of Species by Means of Natural Selection*. John Murray, London.

Davies, N.B., Krebs, J.R. & West, S.A. 2012. *An Introduction to Behavioural Ecology*, 4th edn. Wiley-Blackwell, Oxford, UK.

Dawkins, R. 1979. Twelve misunderstandings of kin selection. *Z. Tierpsychol.* **51**: 184–200.

Dawkins, R. 1982. *The Extended Phenotype*. Freeman, San Francisco.

Dugatkin, L.A. 2007. Inclusive fitness theory from Darwin to Hamilton. *Genetics* **176**: 1375–1380.

Dugatkin, L.A. & Reeve, H.K. 1994. Behavioral ecology and levels of selection: dissolving the group selection controversy. *Adv. Study Behav.* **23**: 101–133.

Eickwort, G.C. 1981. Presocial insects. In: *Social Insects*, Vol. 2 (H.R. Hermann, ed.), pp. 199–280. Academic Press, New York.

Eshel, I. 1972. On the neighbor effect and the evolution of altruistic traits. *Theor. Popul. Biol.* **3**: 258–277.

Evans, H.E. 1977. Extrinsic versus intrinsic factors in the evolution of insect sociality. *Bioscience* **27**: 613–617.

Fawcett, T.W. & Higginson, A.D. 2012. Heavy use of equations impedes communication among biologists. *Proc. Natl Acad. Sci. USA* **109**: 11735–11739.

Fisher, R.A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edinb.* **52**: 399–433.

Fisher, R.A. 1930. *The Genetical Theory of Natural Selection*. Clarendon, Oxford.

Fletcher, J.A. & Doebeli, M. 2009. A simple and general explanation for the evolution of altruism. *Proc. R. Soc. B.* **276**: 13–19.

Foster, K.R., Wenseleers, T. & Ratnieks, F.L.W. 2006. Kin selection is the key to altruism. *Trends Ecol. Evol.* **21**: 57–60.

Frank, S.A. 1983a. A hierarchical view of sex-ratio patterns. *Fla. Entomol.* **66**: 42–75.

Frank, S.A. 1983b. *Theoretical and empirical studies of sex ratios, mainly in fig wasps*. M.S. Thesis, University of Florida, Gainesville.

Frank, S.A. 1985a. Are mating and mate competition by the fig wasp *Pegoscapus assuetus* (Agaonidae) random within a fig? *Biotropica* **17**: 170–172.
Frank, S.A. 1985b. Hierarchical selection theory and sex ratios. II. On applying the theory, and a test with fig wasps. *Evolution* **39**: 949–964.
Frank, S.A. 1986a. Dispersal polymorphisms in subdivided populations. *J. Theor. Biol.* **122**: 303–309.
Frank, S.A. 1986b. The genetic value of sons and daughters. *Heredity* **56**: 351–354.
Frank, S.A. 1986c. Hierarchical selection theory and sex ratios I. General solutions for structured populations. *Theor. Popul. Biol.* **29**: 312–342.
Frank, S.A. 1987. Demography and sex ratio in social spiders. *Evolution* **41**: 1267–1281.
Frank, S.A. 1994. Genetics of mutualism: the evolution of altruism between species. *J. Theor. Biol.* **170**: 393–400.
Frank, S.A. 1995. Mutual policing and repression of competition in the evolution of cooperative groups. *Nature* **377**: 520–522.
Frank, S.A. 1997a. Multivariate analysis of correlated selection and kin selection, with an ESS maximization method. *J. Theor. Biol.* **189**: 307–316.
Frank, S.A. 1997b. The Price equation, Fisher's fundamental theorem, kin selection, and causal analysis. *Evolution* **51**: 1712–1729.
Frank, S.A. 1998. *Foundations of Social Evolution*. Princeton University Press, Princeton, NJ.
Frank, S.A. 2012a. Natural selection. III. Selection versus transmission and the levels of selection. *J. Evol. Biol.* **25**: 227–243.
Frank, S.A. 2012b. Natural selection. IV. The Price equation. *J. Evol. Biol.* **25**: 1002–1019.
Frank, S.A. 2012c. Natural selection. V. How to read the fundamental equations of evolutionary change in terms of information theory. *J. Evol. Biol.* **25**: 2377–2396.
Frank, S.A. 2013. Natural selection. VI. Partitioning the information in fitness and characters by path analysis. *J. Evol. Biol.* **26**: 457–471.
Gardner, A., West, S.A. & Barton, N.H. 2007. The relation between multilocus population genetics and social evolution theory. *Am. Nat.* **169**: 207–226.
Gardner, A., West, S.A. & Wild, G. 2011. The genetical theory of kin selection. *J. Evol. Biol.* **24**: 1020–1043.
Grafen, A. 1984. Natural selection, kin selection and group selection. In: *Behavioural Ecology* (J.R. Krebs, N.B. Davies, eds), pp. 62–84. Blackwell Scientific Publications, Oxford.
Grafen, A. 1985. A geometric view of relatedness. *Oxf. Surv. Evol. Biol.* **2**: 28–89.
Grafen, A. 2006. Optimization of inclusive fitness. *J. Theor. Biol.* **238**: 541–563.
Haldane, J.B.S. 1932. *The Causes of Evolution*. Cornell University Press, Ithaca, NY.
Haldane, J.B.S. 1955. Population genetics. *New Biol.* **18**: 34–51.
Hamilton, W.D. 1963. The evolution of altruistic behavior. *Am. Nat.* **97**: 354–356.
Hamilton, W.D. 1964a. The genetical evolution of social behaviour. I. *J. Theor. Biol.* **7**: 1–16.
Hamilton, W.D. 1964b. The genetical evolution of social behaviour. II. *J. Theor. Biol.* **7**: 17–52.
Hamilton, W.D. 1967. Extraordinary sex ratios. *Science* **156**: 477–488.
Hamilton, W.D. 1970. Selfish and spiteful behaviour in an evolutionary model. *Nature* **228**: 1218–1220.
Hamilton, W.D. 1972. Altruism and related phenomena, mainly in social insects. *Annu. Rev. Ecol. Syst.* **3**: 193–232.
Hamilton, W.D. 1975. Innate social aptitudes of man: an approach from evolutionary genetics. In: *Biosocial Anthropology* (R. Fox, ed.), pp. 133–155. Wiley, New York.
Hamilton, W.D. 1978. Evolution and diversity under bark. In: *Diversity of Insect Faunas* (L.A. Mound, N. Waloff, eds), pp. 154–175. Blackwell Scientific Publications, London.
Hamilton, W.D. 1979. Wingless and fighting males in fig wasps and other insects. In: *Reproductive Competition and Sexual Selection in Insects* (M.S. Blum, N.A. Blum, eds), pp. 167–220. Academic Press, New York.
Hamilton, W.D. 1996. *Narrow Roads of Gene Land*. Freeman, San Francisco.
Hamilton, W.D. & May, R.M. 1977. Dispersal in stable habitats. *Nature* **269**: 578–581.
Heisler, I.L. & Damuth, J. 1987. A method for analyzing selection in hierarchically structured populations. *Am. Nat.* **130**: 582–602.
Henrich, J. 2004. Cultural group selection, coevolutionary processes and large-scale cooperation. *J. Econ. Behav. Organ.* **53**: 3–35.
Karlin, S. & Matessi, C. 1983. Kin selection and altruism. *Proc. R. Soc. B.* **219**: 327–353.
Keller, L. (Ed.). 1999. *Levels of Selection in Evolution*. Princeton University Press, Princeton, NJ.
Kennedy, J.S. 1966. Some outstanding questions in insect behaviour. In: *Symposium of the Royal Entomological Society of London*, Vol. 3 (P.T. Haskell, ed.), pp. 97–112. Royal Entomological Society, London.
Kerr, B., Godfrey-Smith, P. & Feldman, M.W. 2004. What is altruism? *Trends Ecol. Evol.* **19**: 135–140.
Lande, R. & Arnold, S.J. 1983. The measurement of selection on correlated characters. *Evolution* **37**: 1212–1226.
Lazar, M.A. & Birnbaum, M.J. 2012. De-meaning of metabolism. *Science* **336**: 1651–1652.
Lehmann, L. & Keller, L. 2006. The evolution of cooperation and altruism – a general framework and a classification of models. *J. Evol. Biol.* **19**: 1365–1376.
Leigh, E.G. Jr 2010. The group selection controversy. *J. Evol. Biol.* **23**: 6–19.
Levins, R. 1970. Extinction. In: *Some Mathematical Questions in Biology*, Vol. 2 (M. Gerstenhaber, ed.), pp. 77–107. American Mathematical Society, Providence, RI.
Lewontin, R.C. 1970. The units of selection. *Annu. Rev. Ecol. Syst.* **1**: 1–18.
Lin, N. & Michener, C.D. 1972. Evolution of sociality in insects. *Q. Rev. Biol.* **47**: 131–159.
Maynard Smith, J. 1964. Group selection and kin selection. *Nature* **201**: 1145–1147.
Maynard Smith, J. 1971. The origin and maintenance of sex. In: *Group Selection* (G.C. Williams, ed.), pp. 163–175. Aldine-Atherton, Chicago, IL.
Maynard Smith, J. 1976. Group selection. *Q. Rev. Biol.* **51**: 277–283.
Maynard Smith, J. 1982. *Evolution and the Theory of Games*. Cambridge University Press, Cambridge.
Maynard Smith, J. & Szathmáry, E. 1995. *The Major Transitions in Evolution*. Freeman, San Francisco.
McGlothlin, J.W., Moore, A.J., Wolf, J.B. & Brodie, E.D. III 2010. Interacting phenotypes and the evolutionary process. III. Social evolution. *Evolution* **64**: 2558–2574.

Michener, C.D. 1974. *The Social Behavior of Bees: A Comparative Study*. Harvard University Press, Cambridge, MA.
Michener, C.D. & Brothers, D.J. 1974. Were workers of eusocial Hymenoptera initially altruistic or oppressed? *Proc. Natl Acad. Sci. USA* **71**: 671–674.
Michod, R.E. 1982. The theory of kin selection. *Annu. Rev. Ecol. Syst.* **13**: 23–55.
Michod, R.E. & Roze, D. 2001. Cooperation and conflict in the evolution of multicellularity. *Heredity* **86**: 1–7.
Motro, U. 1982a. Optimal rates of dispersal I. Haploid populations. *Theor. Popul. Biol.* **21**: 394–411.
Motro, U. 1982b. Optimal rates of dispersal II. Diploid populations. *Theor. Popul. Biol.* **21**: 412–429.
Motro, U. 1983. Optimal rates of dispersal III. Parent-offspring conflict. *Theor. Popul. Biol.* **23**: 159–168.
Nowak, M.A., Tarnita, C.E. & Wilson, E.O. 2010. The evolution of eusociality. *Nature* **466**: 1057–1062.
Nunney, L. 1985. Group selection, altruism, and structured-deme models. *Am. Nat.* **126**: 212–230.
Okasha, S. 2006. *Evolution and the Levels of Selection*. Oxford University Press, New York.
Pearson, K. 1903. Mathematical contributions to the theory of evolution. XI. On the influence of natural selection on the variability and correlation of organs. *Philos. Trans. R. Soc. Lond. Ser. A* **200**: 1–66.
Pierce, J.R. 1980. *An Introduction to Information Theory: Symbols, Signals and Noise*. Dover Publications, New York.
Price, G.R. 1970. Selection and covariance. *Nature* **227**: 520–521.
Price, G.R. 1972. Extension of covariance selection mathematics. *Ann. Hum. Genet.* **35**: 485–490.
Queller, D.C. 1985. Kinship, reciprocity and synergism in the evolution of social behaviour. *Nature* **308**: 366–367.
Queller, D.C. 1992a. A general model for kin selection. *Evolution* **46**: 376–380.
Queller, D.C. 1992b. Quantitative genetics, inclusive fitness, and group selection. *Am. Nat.* **139**: 540–558.
Queller, D.C. 1994. Genetic relatedness in viscous populations. *Evol. Ecol.* **8**: 70–73.
Queller, D.C. & Strassmann, J.E. 1998. Kin selection and social insects. *Bioscience* **48**: 165–175.
Read, C. 1909. *Logic, Deductive and Inductive*. A. Moring, London.
Rousset, F. 2004. *Genetic Structure and Selection in Subdivided Populations*. Princeton University Press, Princeton, NJ.
Samuelson, P.A. 1983. *Foundations of Economic Analysis*, Enlarged edn. Harvard University Press, Cambridge, MA.
Schumpeter, J.A. 1954. *History of Economic Analysis*. Oxford University Press, New York.
Sober, E. & Wilson, D.S. 1998. *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Harvard University Press, Cambridge, MA.
Soltis, J., Boyd, R. & Richerson, P.J. 1995. Can group-functional behaviors evolve by cultural group selection? An empirical test. *Curr. Anthropol.* **36**: 473–494.
Taylor, P.D. 1992a. Altruism in viscous populations – an inclusive fitness approach. *Evol. Ecol.* **6**: 352–356.
Taylor, P.D. 1992b. Inclusive fitness in a heterogeneous environment. *Proc. R. Soc. B.* **249**: 299–302.
Taylor, P.D. & Frank, S.A. 1996. How to make a kin selection model. *J. Theor. Biol.* **180**: 27–37.
Taylor, P.D., Wild, G. & Gardner, A. 2006. Direct fitness or inclusive fitness: how shall we model kin selection? *J. Evol. Biol.* **20**: 301–309.
Traulsen, A. & Nowak, M.A. 2006. Evolution of cooperation by multilevel selection. *Proc. Natl Acad. Sci. USA* **103**: 10952–10955.
Trivers, R. 1985. *Social Evolution*. Benjamin/Cummings, Menlo Park, CA.
Trivers, R.L. & Hare, H. 1976. Haplodiploidy and the evolution of the social insects. *Science* **191**: 249–263.
Uyenoyama, M. & Feldman, M.W. 1980. Theories of kin and group selection: a population genetics perspective. *Theor. Popul. Biol.* **17**: 380–414.
Uyenoyama, M.K. & Feldman, M. 1982. Population genetic theory of kin selection. II. The multiplicative model. *Am. Nat.* **120**: 614–627.
Uyenoyama, M.K., Feldman, M. & Mueller, L.D. 1981. Population genetic theory of kin selection: multiple alleles at one locus. *Proc. Natl Acad. Sci. USA* **78**: 5036–5040.
Wade, M.J. 1978. A critical review of the models of group selection. *Q. Rev. Biol.* **53**: 101–114.
Wade, M.J. 1985. Soft selection, hard selection, kin selection, and group selection. *Am. Nat.* **125**: 61–73.
Wenseleers, T. 2006. Modelling social evolution: the relative merits and limitations of a Hamilton's rule-based approach. *J. Evol. Biol.* **19**: 1419–1422.
West, S.A. 2009. *Sex Allocation*. Princeton University Press, Princeton, NJ.
West, S.A., Diggle, S.P., Buckling, A., Gardner, A. & Griffin, A.S. 2007. The social lives of microbes. *Annu. Rev. Ecol. Syst.* **38**: 53–77.
West, S.A., Griffin, A.S. & Gardner, A. 2008. Social semantics: how useful has group selection been? *J. Evol. Biol.* **21**: 374–385.
West-Eberhard, M.J. 1975. The evolution of social behavior by kin selection. *Q. Rev. Biol.* **50**: 1–34.
West-Eberhard, M.J. 1976. Born: Sociobiology. *Q. Rev. Biol.* **51**: 89–92.
West-Eberhard, M.J. 1978. Polygyny and the evolution of social behavior in wasps. *J. Kansas Entomol. Soc.* **51**: 832–856.
Williams, G.C. 1966. *Adaptation and Natural Selection*. Princeton University Press, Princeton, NJ.
Williams, G.C. & Williams, D.C. 1957. Natural selection of individually harmful social adaptations among sibs with special reference to social insects. *Evolution* **11**: 32–39.
Wilson, D.S. 1975a. A theory of group selection. *Proc. Natl Acad. Sci. USA* **72**: 143–146.
Wilson, D.S. 1983. The group selection controversy: history and current status. *Annu. Rev. Ecol. Syst.* **14**: 159–187.
Wilson, D.S. & Colwell, R.K. 1981. The evolution of sex ratio in structured demes. *Evolution* **35**: 882–897.
Wilson, D.S., Pollock, G.B. & Dugatkin, L.A. 1992. Can altruism evolve in purely viscous populations? *Evol. Ecol.* **6**: 331–341.
Wilson, E.O. 1971. *The Insect Societies*. Harvard University Press, Cambridge, MA.
Wilson, E.O. 1973. Group selection and its significance for ecology. *Bioscience* **23**: 631–638.
Wilson, E.O. 1975b. The origin of sex. *Science* **188**: 139–140.
Wilson, E.O. 1975c. *Sociobiology: The New Synthesis*. Belknap Press, Harvard, Cambridge, MA.
Wolf, J.B. & Moore, A.J. 2010. Interacting phenotypes and indirect genetic effects: a genetic perspective on the evolu-

tion of social behavior. In: *Evolutionary Behavioural Ecology* (C.W. Fox, D.F. Westneat, eds), pp. 225–245. Oxford University Press, New York.
Wolf, J.B., Brodie, E.D., Cheverud, J.M., Moore, A.J. & Wade, M.J. 1998. Evolutionary consequences of indirect genetic effects. *Trends Ecol. Evol.* **13**: 64–69.
Wolf, J.B., Brodie, E.D. III & Moore, A.J. 1999. Interacting phenotypes and the evolutionary process. II. Selection resulting from social interactions. *Am. Nat.* **153**: 254–266.
Wright, S. 1945. Tempo and mode in evolution: a critical review. *Ecology* **26**: 415–419.

Supporting information

Additional Supporting Information may be found in the online version of this article:
Appendix S1 Notes distributed by W.D. Hamilton for his graduate class, Fall 1979.
Appendix S2 Notes distributed by W.D. Hamilton for his graduate class, Fall 1979.

Received 27 September 2012; revised 26 November 2012; accepted 29 November 2012