Published in final edited form as:
J Evol Biol. 2012 February ; 25(2): 227–243. doi:10.1111/j.1420-9101.2011.02431.x.

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

Steven A. Frank†
Steven A. Frank: safrank@uci.edu
†Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697–2525, USA

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.

Keywords

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

> 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.

*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.

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 and Demeter's (1987) stochastic corrector model of early cellular evolution, Levin and Bull's (1994) short-sighted model of parasite evolution, Frank's (2010) timescale model of microbial metabolism, and Maynard Smith and 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.

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 neighbors (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 favored 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 of the partition and the consequences for understanding evolutionary process (Box 2).

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 Eq. (1) in the text. The Price equation may be written as

$$\overline{w}\Delta\overline{z} = \text{Cov}(w, z) + \text{E}(w\Delta z).$$

Comparing to Eq. (1), the selection bias is $\Delta S = \text{Cov}(w, z)$. 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 $\Delta\tau = \text{E}(w\Delta z)$. 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 Eq. (4) is derived as $\Delta S = \text{Cov}(w, z) = \beta_{wz}V_z$, 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 Eq. (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

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

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, Bürger, Butcher & Gabriel, 1993).

Note that the regression expression $E(w|z) = 1 - \beta_{wz} z$ 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.

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

$$\bar{w}\Delta\bar{z}=\Delta S+\Delta\tau. \quad (1)$$

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

$$\bar{w}\Delta\bar{z}=\Delta S+\Delta\tau=0, \quad (2)$$

which also means that at an equilibrium balance

$$\Delta S=-\Delta\tau. \quad (3)$$

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

$$\Delta S=s_z V_z, \quad (4)$$

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 Eq. (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 *sz* and *Vz* 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, $\mathcal{R}$, between the forces. In particular, when the two terms oppose each other, we may write

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

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 Eq. (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 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, however 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 Eq. (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 Eq. (3), the balance occurs when $\Delta S = -\Delta\tau$. From Eq. (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 Eq. (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 Eq. (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 Eq. (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 versus 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, $\mathcal{R}$, from Eq. (7). For example, Fig. 1a shows the first model with equilibrium $q = \mu/s$ in Eq. (9). In that case, $\Delta S = -sq(1-q)$ and $\Delta\tau = \mu(1-q)$, so the ratio is $\mathcal{R} = \log(sq/\mu)$.

The top curve of Fig. 1a plots $\mathcal{R}$ 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 and low frequencies 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, $\mathcal{R} \gg 0$. 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 $\mathcal{R} = 0$, at which point the selection bias equals the transmission bias. Similarly, when q is very small, then the transmission bias is much greater than the selection bias, $\mathcal{R} \ll 0$, and the mutant frequency increases until the point $\mathcal{R} = 0$.

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 of selection and transmission, how should we interpret Williams' dictum?

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 toward 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 and 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 Eqs. (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 Eq. (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 Eq. (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 Eq. (9) from the previous section. Box 4 provides the derivation.

Box 4

Variance under mutation-selection balance

To obtain the equilibrium genetic variance in Eq. (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 Eq. (12) leads to Eq. (13).

To obtain the equilibrium genetic variance in Eq. (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), 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 Eq. (12) yields Eq. (14).

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

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

which matches the result for the diploid recessive model in Eq. (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 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 Eq. (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 magnitude of selection, $s \ll \mu$, which corresponds to results above for weak selection.

From Eq. (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 Eq. (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, McCaskill and Schuster (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 (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 favors 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

favoring sexuals, with rate s, against transmission bias within clades favoring 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 Eq. (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 toward larger body size. The distribution of mammalian body size is influenced by the balance between selection between genera favoring smaller size and selection within genera favoring 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 lack of fidelity in transmission (Eq. 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 Eq. (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 regulates 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, Pollock & Dugatkin, 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, then 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 $\Delta S_\alpha = -\Delta S_w$, 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 $V_\mathrm{w} = V_\mathrm{t} - V_\mathrm{a}$ yields

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

It is convenient to express the pattern of variance by the correlation coefficient $r = V_\alpha/V_t$, 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, $r = V_\alpha/V_t$. 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, $s_\alpha r$ is the intensity of selection among groups, s_α, 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 analyzed more formally (Grafen, 1984; Wade, 1985; Frank, 1986, 1998; West, Griffin & Gardner, 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 neighboring group members for access to local resources, and z_g is the average competitive tendency in the individual's group. Selection within groups always favors 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 neighbors, the less efficient the group is in using its resources productively. For example, a certain fraction of local energy may go into outcompeting neighbors 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 favors a less competitive and more cooperative behavioral 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 Eq. (15). The selective intensity among groups is $s_\alpha = -1$, 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_α and s_w into Eq. (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 toward a single value that is an evolutionarily stable strategy (ESS). Thus, individual values, z, converge toward group averages, z_g, which in turn converge to a global value, z^*. Making the substitution $z_g = z^*$ and solving for z^* gives the balance point (Frank, 1994, 1995b) as

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

This balance point expresses the key insights of 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 neighbors. As variance shifts toward the group level, $1-r$ declines, z^* decreases because competitive restraint is more strongly favored, and the balance of selective forces increasingly favors cooperative behavior. In terms of kin selection, as the coefficient of relatedness, r, increases, competitive restraint and cooperative behavior 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 regulates 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 $s_\alpha = -1$. 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 $1 - \varepsilon z_g$, where $\varepsilon < 1$, then $s_\alpha = -\varepsilon$. Using that value of s_α in Eq. (17), we obtain the solution

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

This solution is equivalent to Eq. (18) when $\varepsilon = 1$. As limits on group productivity become more stringent, ε declines toward zero, the balance tips more strongly in favor 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 Eq. (7) compares selection among groups to selection within groups. Substituting the expressions for ΔS and $\Delta\tau$ derived from Eq. (16) into Eq. (7) yields

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

Fig. 2 plots $\mathcal{R}$ 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 $\mathcal{R} = 0$. 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 toward the higher level.

That interpretation of group-level unity or adaptation goes beyond what the analysis by itself presents. The analysis simply describes the way in which $\mathcal{R}$ 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 Eq. (16), repeated here

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

where z is a chromosome's tendency to be competitive against neighboring chromosomes for access to local resources within the protocell, and z_g is the average competitive tendency of chromosomes in the protocell. Following Eq. (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 favored to reduce the number, k, of symbionts transmitted to offspring or transmitted between hosts (Frank, 1996a). Hosts may also be favored 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 favors 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 causes a transmission bias that favors 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 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 among mutation, selection and segregation (Frank, 1994). The observed equilibrium trait values in the computer simulation closely follow the prediction $z^* = 1 - r$, where $r = V_\alpha/V_t$ 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 patterns of variance and correlation, not on genealogy (Frank, 1998). Genealogy is often closely associated with 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 favor kin selection over group selection

Kin 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 favored value of competitiveness was given in Eq. (18) by $z^* = 1 - r$. Competitiveness rises as relatedness between group members, r, declines. The equilibrium fitness from Eq. (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 and 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 favor 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 toward characters favored within groups. For example, within-group selection often favors greater competitiveness against neighbors. 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 neighbors 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, Schuster & Bonhoeffer, 2001).

Once again, we have a situation in which selection within groups favors more competitive traits, whereas selection between groups favors 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 rate of resource uptake. Individuals that invest more energy in uptake outcompete neighbors 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 setup 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

Fig. 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 Eq. (1) that the total change in a character is

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

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. 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 favored to grow quickly (high rate and low efficiency) to use up available resources before extinction, whereas long-lived colonies are favored to grow slowly and use resources efficiently. Thus, in the lower curve of Fig. 3, longer colony survival causes the optimal balance to shift toward lower rate and higher yield.

Balance between selection and transmission

When mutation generates variation within colonies, then the rate-yield tradeoff 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 toward greater investment in resource acquisition. Higher resource acquisition and metabolic rate arise from the inevitable production of mutant neighbors 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 tradeoff 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 Eq. (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, $\mathcal{R}$. 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, then the relative dominance of one level over the other often sets the level at which functional coherence and individuality emerges.

A key aspect of Maynard Smith and Szathmáry's (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 expresses 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 analyze 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 RD, Borgia G. Group selection, altruism, and the levels of organization of life. Annual Review of Ecology and Systematics. 1978; 9:449–474.

Barton NH, Turelli M. Adaptive landscapes, genetic distance and the evolution of quantitative characters. Genetical Research. 1987; 49:157–173. [PubMed: 3596236]

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

Eigen M. Self-organization of matter and the evolution of biological macromolecules. Naturwissenschaften. 1971; 58:465–523. [PubMed: 4942363]

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

Eigen M, McCaskill J, Schuster P. Molecular quasi-species. Journal of Physical Chemistry. 1988; 92:6881–6891.

Eigen M, McCaskill J, Schuster P. The molecular quasi-species. Advances in Chemical Physics. 1989:149–263.

Eigen M, Schuster P. The hypercycle. A principle of natural self-organization. Part A: emergence of the hypercycle. Naturwissenschaften. 1977; 64:541–565. [PubMed: 593400]

Fisher, RA. The Genetical Theory of Natural Selection. Clarendon; Oxford: 1930.

Frank SA. Hierarchical selection theory and sex ratios I. General solutions for structured populations. Theoretical Population Biology. 1986; 29:312–342. [PubMed: 3738836]

Frank SA. Demography and sex ratio in social spiders. Evolution. 1987; 41:1267–1281.

Frank SA. Kin selection and virulence in the evolution of protocells and parasites. Proceedings of the Royal Society of London B. 1994; 258:153–161.

Frank SA. George Price's contributions to evolutionary genetics. Journal of Theoretical Biology. 1995a; 175:373–388. [PubMed: 7475081]

Frank SA. Mutual policing and repression of competition in the evolution of cooperative groups. Nature. 1995b; 377:520–522. [PubMed: 7566147]

Frank SA. Host control of symbiont transmission: the separation of symbionts into germ and soma. American Naturalist. 1996a; 148:1113–1124.

Frank SA. Host-symbiont conflict over the mixing of symbiotic lineages. Proceedings of the Royal Society of London B. 1996b; 263:339–344.

Frank SA. Models of parasite virulence. Quarterly Review of Biology. 1996c; 71:37–78. [PubMed: 8919665]

Frank SA. Models of symbiosis. American Naturalist. 1997a; 150:S80–S99.

Frank SA. The Price equation, Fisher's fundamental theorem, kin selection, and causal analysis. Evolution. 1997b; 51:1712–1729.

Frank, SA. Foundations of Social Evolution. Princeton University Press; Princeton, New Jersey: 1998.

Frank SA. The trade-off between rate and yield in the design of microbial metabolism. Journal of Evolutionary Biology. 2010; 23:609–613. [PubMed: 20487133]

Frank SA, Slatkin M. The distribution of allelic effects under mutation and selection. Genetical Research. 1990; 55:111–117. [PubMed: 2370007]

Gardner A. The Price equation. Current Biology. 2008; 18:R198–R202. [PubMed: 18334191]

Gardner A, Grafen A. Capturing the superorganism: a formal theory of group adaptation. Journal of Evolutionary Biology. 2009; 22:659–671. [PubMed: 19210588]

Grafen, A. Natural selection, kin selection and group selection. In: Krebs, JR.; Davies, NB., editors. Behavioural ecology. Blackwell Scientific Publications; Oxford: 1984. p. 62-84.

Grafen A. A first formal link between the price equation and an optimization program. Journal of Theoretical Biology. 2002; 217:75–91. [PubMed: 12183132]

Haldane JBS. A mathematical theory of natural and artificial selection, Part V: Selection and mutation. Mathematical Proceedings of the Cambridge Philosophical Society. 1927; 23:838–844.

Hamilton WD. Selfish and spiteful behaviour in an evolutionary model. Nature. 1970; 228:1218–1220. [PubMed: 4395095]

Hamilton, WD. Innate social aptitudes of man: an approach from evolutionary genetics. In: Fox, R., editor. Biosocial Anthropology. Wiley; New York: 1975. p. 133-155.

Hamilton, WD. Wingless and fighting males in fig wasps and other insects. In: Blum, MS.; Blum, NA., editors. Reproductive Competition and Sexual Selection in Insects. Academic Press; New York: 1979. p. 167-220.

Heisler IL, Damuth J. A method for analyzing selection in hierarchically structured populations. American Naturalist. 1987; 130:582–602.

Kimura M. A stochastic model concerning the maintenance of genetic variability in quantitative characters. Proceedings of the National Academy of Sciences USA. 1965; 54:731–736.

Lande R. The maintenance of genetic variability by mutation in a polygenic character with linked loci. Genetical Research. 1975; 26:221–235. [PubMed: 1225762]

Levin BR, Bull JJ. Short-sighted evolution and the virulence of pathogenic microorganisms. Trends in Microbiology. 1994; 2:76–81. [PubMed: 8156275]

Lewontin RC. The units of selection. Annual Review of Ecology and Systematics. 1970; 1:1–18.

Lynch M, Bürger R, Butcher D, Gabriel W. The mutational meltdown in asexual populations. Journal of Heredity. 1993; 84:339–344. [PubMed: 8409355]

Maynard Smith J. Group selection. Quarterly Review of Biology. 1976; 51:277–283.

Maynard Smith J. Hypercycles and the origin of life. Nature. 1979; 280:445–446. [PubMed: 460422]

Maynard Smith, J. Evolutionary progress and levels of selection. In: Nitecki, MH., editor. Evolutionary Progress. University of Chicago Press; Chicago: 1988. p. 219-230.

Maynard Smith, J.; Szathmáry, E. The Major Transitions in Evolution. Freeman; San Francisco: 1995.

Michod RE. Cooperation and conflict in the evolution of individuality. I. Multilevel selection of the organism. American Naturalist. 1997a; 149:607–645.

Michod RE. Evolution of the individual. The American Naturalist. 1997b; 150:S5–S21.

Michod RE. The group covariance effect and fitness trade-offs during evolutionary transitions in individuality. Proceedings of the National Academy of Sciences. 2006; 103:9113–9117.

Michod RE, Herron MD. Cooperation and conflict during evolutionary transitions in individuality. Journal of evolutionary biology. 2006; 19:1406–1409. [PubMed: 16910968]

Michod RE, Nedelcu AM. On the reorganization of fitness during evolutionary transitions in individuality. Integrative and Comparative Biology. 2003; 43:64–73. [PubMed: 21680410]

Nachman M, Crowell S. Estimate of the mutation rate per nucleotide in humans. Genetics. 2000; 156:297–304. [PubMed: 10978293]

Okasha, S. Evolution and the Levels of Selection. Oxford University Press; New York: 2006.

Pfeiffer T, Schuster S, Bonhoeffer S. Cooperation and competition in the evolution of ATP-producing pathways. Science. 2001; 292:504–7. [PubMed: 11283355]

Price GR. Selection and covariance. Nature. 1970; 227:520–521. [PubMed: 5428476]

Price GR. Extension of covariance selection mathematics. Annals of Human Genetics. 1972; 35:485–490. [PubMed: 5073694]

Queller DC. Genetic relatedness in viscous populations. Evolutionary Ecology. 1994; 8:70–73.

Rice, SH. Evolutionary Theory: Mathematical and Conceptual Foundations. Sinauer Associates; Sunderland, MA: 2004.

Sober E, Wilson DS. A critical review of philosophical work on the units of selection problem. Philosophy of Science. 1994:534–555.

Stearns, SC. Natural selection and fitness, adaptation and constraint. In: Jablonski, D.; Raup, DM., editors. Patterns and Processes in the History of Life. Springer-Verlag; Berlin: 1986. p. 23-44.

Stebbins, GL. Variation and Evolution in Plants. Columbia University Press; New York: 1950.

Szathmáry E, Demeter L. Group selection of early replicators and the origin of life. Journal of Theoretical Biology. 1987; 128:463–486. [PubMed: 2451771]

Taylor PD. Altruism in viscous populations|an inclusive fitness approach. Evolutionary Ecology. 1992; 6:352–356.

Turelli M. Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. Theoretical Population Biology. 1984; 25:138–193. [PubMed: 6729751]

Van Valen LM. Group selection, sex, and fossils. Evolution. 1975; 29:87–94.

Van Valen LM. Species, sets, and the derivative nature of philosophy. Biology and Philosophy. 1988; 3:49–66.

Wade MJ. Soft selection, hard selection, kin selection, and group selection. American Naturalist. 1985; 125:61–73.

West SA, Griffin AS, Gardner A. Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. Journal of Evolutionary Biology. 2007; 20:415–432. [PubMed: 17305808]

Wilke C. Quasispecies theory in the context of population genetics. BMC Evolutionary Biology. 2005; 5:44. [PubMed: 16107214]

Williams, GC. Adaptation and Natural Selection. Princeton University Press; Princeton, NJ: 1966.

Williams, GC. Natural Selection: Domains, Levels, and Challenges. Oxford University Press; Oxford: 1992.

Wilson DS, Colwell RK. The evolution of sex ratio in structured demes. Evolution. 1981; 35:882–897.

Wilson DS, Pollock GB, Dugatkin LA. Can altruism evolve in purely viscous populations? Evolutionary Ecology. 1992; 6:331–341.

Wilson DS, Sober E. Reviving the superorganism. Journal of Theoretical Biology. 1989; 136:337–356. [PubMed: 2811397]

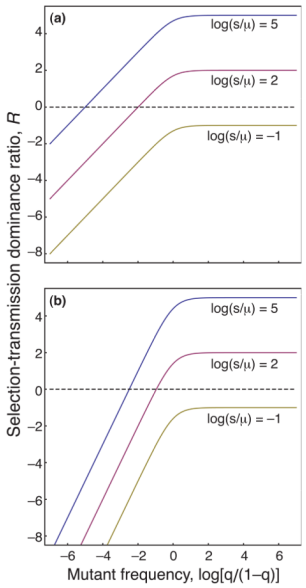

Fig. 1.

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

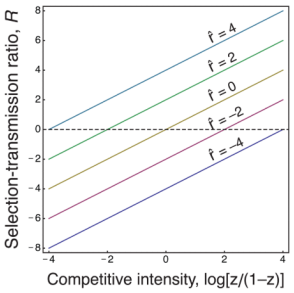

Fig. 2.

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, $\mathcal{R}$, of Eq. (19). Different levels of relatedness shift the balance between selection bias among groups and transmission bias within groups. Here, relatedness is measured by $\hat{r} = \log[r/(1 - r)]$.

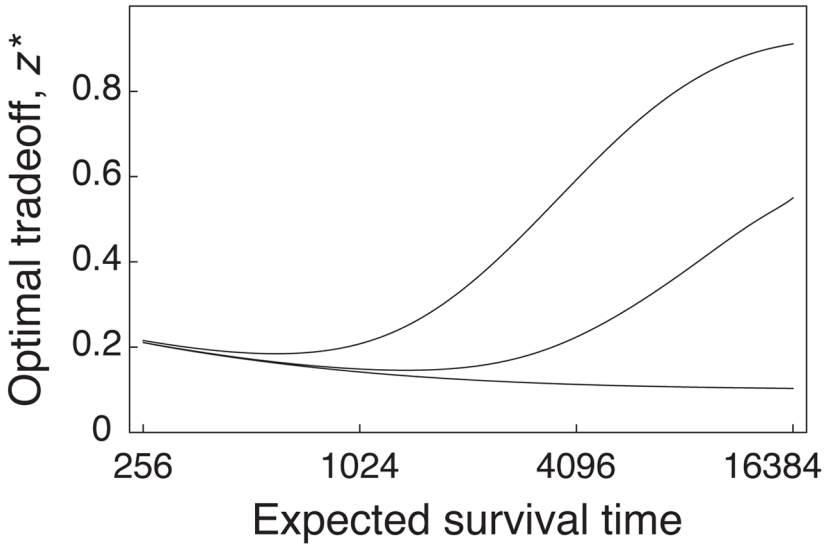

Fig. 3.

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 10μ. 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).