Theoretical Population Biology 104 (2015) 59–67

Contents lists available at ScienceDirect

Theoretical Population Biology

journal homepage: www.elsevier.com/locate/tpb

On the interpretation and relevance of the Fundamental Theorem of Natural Selection

Warren J. Ewens [a,*], Sabin Lessard [b]

[a] Department of Biology, The University of Pennsylvania, Philadelpia, PA 19104, USA

[b] Département de mathématiques et de statistique, Université de Montréal, Montréal, Québec H3C 3J7, Canada

ARTICLE INFO

Article history:
Received 18 February 2015
Available online 26 July 2015

Keywords:
Average effects
Average excesses
Fitness
Gene frequencies
Genic evolutionary potential
Mean fitness

ABSTRACT

The attempt to understand the statement, and then to find the interpretation, of Fisher's "Fundamental Theorem of Natural Selection" caused problems for generations of population geneticists. Price's (1972) paper was the first to lead to an understanding of the statement of the theorem. The theorem shows (in the discrete-time case) that the so-called "partial change" in mean fitness of a population between a parental generation and an offspring generation is the parental generation additive genetic variance in fitness divided by the parental generation mean fitness. In the continuous-time case the partial rate of change in mean fitness is equal to the parental generation additive genetic variance in fitness with no division by the mean fitness. This "partial change" has been interpreted by some as the change in mean fitness due to changes in gene frequency, and by others as the change in mean fitness due to natural selection. (Fisher variously used both interpretations.) In this paper we discuss these interpretations of the theorem. We indicate why we are unhappy with both. We also discuss the long-term relevance of the Fundamental Theorem of Natural Selection, again reaching a negative assessment. We introduce and discuss the concept of genic evolutionary potential. We finally review an optimizing theorem that involves changes in gene frequency, the additive genetic variance in fitness and the mean fitness itself, all of which are involved in the Fundamental Theorem of Natural Selection, and which is free of the difficulties in interpretation of the Fundamental Theorem of Natural Selection.

1. Introduction

The Fundamental Theorem of Natural Selection (FTNS) as now understood is very general. It is a "whole genome" result in the sense that in the theorem the fitness of any individual is allowed to depend, in an unspecified way, on its autosomal genomic make-up. (The sex chromosomes are also ignored in the FTNS.) It also holds for any form of mating, random or otherwise. Since the theorem concerns the effect of natural selection only, no mutations are allowed. Various simplifications are made in the statement of the theorem. For example, despite the fact that the FTNS relates to diploid populations, no account is taken in the theorem as presented by Fisher of the existence of two sexes, with possibly different fitness values for any genotype between the two sexes. Next, in effect an infinitely large population is assumed so that no random changes in gene frequency are allowed. Various other simplifications are made, no doubt of necessity, and a completely general theorem awaits development. The same simplifying assumptions that are made in the FTNS will also be made in this paper.

There are various versions of the FTNS. Different versions make different assumptions about "time" and about "fitness". Concerning "time", there are both discrete-time and continuous-time versions of the FTNS, and within these two there are various subdivisions, some still under investigation. Concerning fitness, in the general literature the fitness of any individual is sometimes taken as a parameter determined by that individual's genotype, and sometimes (for example Price, 1970) as the actual number of offspring that individual has (or perhaps half that number in the diploid case). Since the FTNS is a deterministic result in which the population size is taken as infinite, in the discussion of the FTNS in discrete time with non-overlapping generations there is no effective difference between these two definitions, and to be concrete we here use the parameter definition of fitness in this context. Moreover, we assume that for each genotype this parameter does not depend on time and population state.

Fisher's (1958) statement of the FTNS, namely "The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time" is nowadays understood

* Corresponding author.
E-mail address: wewens@sas.upenn.edu (W.J. Ewens).

http://dx.doi.org/10.1016/j.tpb.2015.07.002
0040-5809/

as stating that "The partial rate of change in the mean fitness of any population at any time is equal to the additive genetic variance in fitness in that population at that time". The meaning of the words "partial change " is discussed later, and it is plausible that by using the word "increase" rather than "change" Fisher implied by "increase" the same meaning as "partial change".

Fisher's statement clearly refers to a continuous-time process. For simplicity we consider instead a discrete-generation process in which, at discrete time points 1, 2, 3, . . . , a parental generation gives birth to an offspring generation and then dies. However, the conclusions reached can be extended to models with overlapping generations in continuous time as well as discrete time (see, e.g., Ewens, 1989, Lessard, 1997 and Grafen, 2015). Unless indicated otherwise when discussing this discrete-time case, all measurements are taken at the time of conception of any generation and the word "change" means the change in the quantity considered between parental and offspring generations values at their respective times of conception.

2. Background

Despite previous discussion on the total change in mean fitness (see, e.g., Kempthorne, 1957 and Kimura, 1958), the interpretation of the FTNS remained largely unclear until Price's (1972) analysis. Price's analysis was, at least initially, presented in discrete time although, because he was attempting to validate Fisher's continuous-time result given above, his wording had a continuous-time flavor. Price's analysis is based on a partial change which he denoted $\partial_{NS}\bar{w}$ and which is defined in Eq. (11) in a context of discrete non-overlapping generations. (We have changed Price's M to $\bar{w}$, the modern notation for a mean fitness.) Price's equation (2.7) states that $\partial_{NS}\bar{w}$ is a change in mean fitness between discrete time points and his concluding statement of the theorem (his Eq. (5.14)) involves this change. Thus it can be argued that Price's analysis, despite much of his wording, relates to a discrete-time process. However in melding his result with Fisher's verbal presentation he does not give the discrete-time result presented in Eq. (13).

As discussed below the FTNS is quintessentially a diploid population result, and the entire analysis of the FTNS given below assumes a diploid population. The FTNS depends, in the diploid case, on the concept of the average effect of an allele on fitness, and it is convenient to discuss this concept briefly. There are in fact two average effect concepts, and Fisher (1958, p. 35) confusingly uses both without any clarification. What has been called the α definition (Ewens, 2010) is that the average effects are found by minimizing

$$\sum_g P_g \left\{ w_g - \bar{w} - \left(\sum_{i,j \text{ in } g} \alpha_{ij} \right) \right\}^2 \tag{1}$$

with respect to the α_{ij} values. The inner sum in (1) is sometimes called the "breeding value" of genotype g, and for convenience we denote it by α_g. In the above expression g indexes whole genome genotypes, P_g is the frequency of whole-genome genotype g and w_g is the fitness of this genotype, that is its viability from the time of conception to the time of reproduction, which is assumed to be time- and frequency-independent. Moreover,

$$\bar{w} = \sum_g P_g w_g \tag{2}$$

is the population mean fitness (henceforth simply "mean fitness") and α_{ij} is the "α" version average effect of allele j at gene locus i. In the expression (1) the inner sum is taken over all alleles at all loci in genome, with, for each whole-genome genotype g, α_{ij} occurring once, twice or not at all in the inner sum depending on whether allele j at gene locus i occurs once, twice or not at all in that genotype. The outer sum is then taken over all whole-genome genotypes. These average effects are also required to satisfy the constraint that for each gene locus i,

$$\sum_j \alpha_{ij} p_{ij} = 0, \tag{3}$$

where p_{ij} is the frequency of allele j at gene locus i. (Thus at gene locus "i" the sum $\sum_j p_{ij}$ of the frequencies of the various possible alleles is 1.)

On the other hand, the β definition average effects (Ewens, 2010) are found by minimizing

$$\sum_g P_g \left\{ w_g - \left(\sum_{i,j \text{ in } g} \beta_{ij} \right) \right\}^2 \tag{4}$$

with respect to the β_{ij} values. The inner sum in (4) is sometimes called the "additive genetic value" of genotype g, and for convenience we denote it by β_g. The symbols and summations in this expression have the same meaning as in (1). The additive genetic value of genotype g can be interpreted as the fitness of g as best predicted from the alleles in g counted as many times as they occur. The relation between the α_{ij} values and the β_{ij} values is that for each allele j at each locus i,

$$\beta_{ij} = \alpha_{ij} + \frac{\bar{w}}{2K}, \tag{5}$$

where K is the number of loci in the genome that influence fitness. (From now on, only those loci are considered. The factor 2 arises because all loci are diploid.) There is a constraint on the β_{ij} values, namely

$$\sum_j \beta_{ij} p_{ij} = \frac{\bar{w}}{2K} \tag{6}$$

for each gene locus i, which derives from the constraint (3) on the α_{ij} values.

It is in general very difficult to find explicit expressions for average effects, but fortunately explicit expressions are not needed for the discussion in this paper. For some purposes one definition of average effects is more convenient and for other purposes the other definition is more convenient. Under both definitions the average effects of the various alleles usually depend on their frequencies (an exception will be given in Example 1 below) and thus in general average effects will change from one generation to the next as these frequencies change. This is a central observation when discussing the FTNS.

The whole genome "additive genetic variance" in fitness, denoted here and throughout by σ_A^2, is the sum of squares removed by fitting the α parameters in (1) or, equivalently, the β parameters in (4). Under the constraint (3), we have

$$\sigma_A^2 = \sum_g P_g \alpha_g^2 = \sum_g P_g \left(\sum_{i,j \text{ in } g} \alpha_{ij} \right)^2. \tag{7}$$

It is not easy to write down an explicit formula for σ_A^2, but (as with the average effects) an explicit expression is not necessary for the purposes of this paper.

The population mean fitness $\bar{w}$ in some given parental generation at its time of conception can be expressed via the β definition of average effects instead of the actual genotypic fitnesses in the two equivalent forms

$$\bar{w} = \sum_g P_g \beta_g = 2 \sum_i \sum_j p_{ij} \beta_{ij}, \tag{8}$$

with β_{ij} in the inner sum of the right-hand expression occurring once, twice or not at all depending on whether allele j at gene locus i occurs once, twice or not at all in genotype g. In this expression P_g is the parental generation frequency of the whole-genome genotype g and the "β" average effects are the parental generation values. The factor 2 in the right-hand expression arises again from the diploid nature of the individuals in the population.

The mean fitness of the offspring generation at its time of conception can be written as

$$\bar{w}' = \sum_g P'_g \, \beta'_g = 2 \sum_i \sum_j p'_{ij} \beta'_{ij}, \tag{9}$$

where the prime symbol ($'$) indicates an evaluation in the offspring generation at its time of conception. Note that the average effects β'_{ij} will usually differ from the parental generation values β_{ij}, since the average effects are determined by a least-squares procedure involving the current genotype frequencies. (A similar comment applies for the α_{ij}.) The total inter-generational change in mean fitness is clearly given by

$$\bar{w}' - \bar{w} = \sum_g (P'_g \, \beta'_g - P_g \, \beta_g) = 2 \sum_i \sum_j (p'_{ij} \beta'_{ij} - p_{ij} \beta_{ij}). \tag{10}$$

As stated above, parental and offspring generation calculations are taken at their respective times of conception, and "change" refers to the change between these two times.

Price's (1972) natural selection component $\partial_{NS} \bar{w}$ of the inter-generational change in mean fitness is defined by

$$\partial_{NS} \bar{w} = \sum_g (P'_g - P_g) \beta_g = 2 \sum_i \sum_j (p'_{ij} - p_{ij}) \beta_{ij}. \tag{11}$$

The suffix "NS" stands for "natural selection". This change is obtained by keeping the additive genetic values fixed from one generation to the next. It is clear that $\partial_{NS} \bar{w}$ differs from the total change in mean fitness by an amount

$$\partial_{EC} \bar{w} = \sum_g (P'_g \beta'_g - P'_g \beta_g) = 2 \sum_i \sum_j (p'_{ij} \beta'_{ij} - p'_{ij} \beta_{ij}). \tag{12}$$

Price called $\partial_{EC} \bar{w}$ the change in mean fitness due to "environment change", abbreviated EC. This component of the total change includes changes due to epistasis, dominance, mating and recombination. This is a quite dubious interpretation of the word "environment", about which Price himself expresses severe reservations, as discussed below. Gene interactions and rearrangement of genes within individuals because of mating and recombination also influence gene frequency changes and should not be ignored, as also discussed below.

The right-hand side in (11) has been called a "partial change" in mean fitness and shown to be equal to $\sigma_A^2 / \bar{w}$ (Ewens, 1989), where σ_A^2 is the parental generation whole-genome additive genetic variance and $\bar{w}$ is the mean fitness. This leads to the correct statement of the FTNS in the discrete-generation case, in Price's notation, as

$$\partial_{NS} \bar{w} = \frac{\sigma_A^2}{\bar{w}}. \tag{13}$$

As stated above, Price's (1972) Eq. (5.14), corresponding to Eq. (13), does not have the mean fitness $\bar{w}$ in the denominator of the right-hand side, since his aim was to prove Fisher's continuous-time statement of the theorem. The first appearance in the literature of the discrete-time Eq. (13) is in Ewens (1989). We return to further consideration of the deceptively simple Eq. (13) later.

There is another expression of the partial change in mean fitness that is suggested by Fisher's (1941) explanations of the FTNS. The total inter-generational change in mean fitness is obviously

$$\bar{w}' - \bar{w} = \sum_g (P'_g - P_g) w_g = \sum_g (\Delta P_g) w_g, \tag{14}$$

where $\Delta P_g = P'_g - P_g$ is the change in the frequency of genotype g from the parental generation to the offspring generation both at their time of conception. The partial change takes the form (Lessard, 1997)

$$\partial_{NS} \bar{w} = \sum_g \left[\Delta P_g\right]_\alpha w_g, \tag{15}$$

where

$$\left[\Delta P_g\right]_\alpha = \frac{P_g \alpha_g}{\bar{w}}, \tag{16}$$

with α_g being the breeding value of genotype g. The quantity $\left[\Delta P_g\right]_\alpha$ is interpreted as the change in the frequency of genotype g as best predicted from the changes in gene frequencies. Therefore, the same interpretation holds for the partial change in mean fitness.

Although Fisher (1958, pp. 22–30) discussed life-table matters immediately before his presentation of the FTNS, there is no explicit use of these tables in his statement of the theorem. (Fisher makes a passing reference (1958, p. 37) in his development of the FTNS to m, the Malthusian parameter of its life-table analysis, but there is no explicit link in the mathematical aspects of the FTNS development to the life-table work.) Rigorous treatments of the FTNS in age-structured populations with genotype- and age-specific fitness parameters and reproductive values can be found in Crow (1979) for one-locus models and Lessard (1997) for multilocus models. Grafen (2015) gives a version of the FTNS based on individual birth and death rates and population age-specific reproductive values defining fitness, instead of genotypic Malthusian parameters. This definition corresponds to an assumption of demographic equilibrium as used in Charlesworth (1994) based on approximations under weak selection, but was not used in Price (1972).

3. Some simple examples

In order to illustrate the comments made later it is convenient to consider some simple examples of evolutionary calculations. The examples in this section consider three different fitness arrays and (except in Example 3) two mating schemes, random mating and selfing. In all three examples the fitness of any individual depends on its genotype at a single locus A admitting two alleles, A_1 and A_2. The frequencies of the genotypes A_1A_1, A_1A_2 and A_2A_2 at the time of conception are denoted by P_{11}, $2P_{12}$ and P_{22}, respectively, and their fitnesses by w_{11}, w_{12} and w_{22}, respectively. Also, since only one gene locus is involved we use a simpler notation and write the frequency of the allele A_1 as $p = P_{11} + P_{12}$. Note that the mean fitness can then be written as

$$\bar{w} = 2\{p\beta_1 + (1-p)\beta_2\}. \tag{17}$$

where β_j is the "β" version average effect of allele A_j ($j = 1, 2$).

In all three examples the frequencies of the genotypes A_1A_1, A_1A_2 and A_2A_2 at the time of conception of generation 1 are assumed first to be, respectively, 0.2, 0.4 and 0.4. These are not in Hardy–Weinberg form, so that the previous generation did not mate at random. The numerical values given for the next generations refer to their values at their respective times of conception. The results for 3 consecutive generations are given in Tables 1–3, respectively, for Examples 1–3. Any numerical value in those tables which is given to six decimal places is an approximation accurate to that level.

Example 1. In this example the fitnesses of the genotypes A_1A_1, A_1A_2 and A_2A_2 are assumed to be 1.1, 1.0 and 0.9 respectively, so that there is no dominance in the fitnesses at this locus, implying that throughout the evolutionary process $\sigma_A^2 = \sigma^2$. Thus the values of σ^2 are not listed in Table 1. Also, since the β version average

Table 1
Values of various quantities in Example 1 with 1.1, 1.0 and 0.9 as fitness values and 0.2, 0.4 and 0.4 as initial frequencies of genotypes A_1A_1, A_1A_2 and A_2A_2, respectively.

Generation	p	$\bar{w}$	α_1	α_2	σ_A^2	$\sigma_A^2/\bar{w}$
1	0.400000	0.980000	0.060000	−0.040000	0.005600	0.011429
2 (random)	0.428571	0.985714	0.057143	−0.428570	0.004898	0.004969
2 (selfing)	0.428571	0.985714	0.057143	−0.428570	0.007755	0.007868
3 (random)	0.453416	0.990683	0.054658	−0.045342	0.004957	0.005004
3 (selfing)	0.467909	0.993582	0.051661	−0.048339	0.008924	0.008981

Table 2
Values of various quantities in Example 2 with 1.2, 1.0 and 0.9 as fitness values and 0.2, 0.4 and 0.4 as initial frequencies of genotypes A_1A_1, A_1A_2 and A_2A_2, respectively.

Generation	p	$\bar{w}$	β_1	β_2	σ^2	σ_A^2	$\sigma_A^2/\bar{w}$	σ_A^2/σ^2
1	0.400000	1.000000	0.585714	0.442857	0.012000	0.011429	0.011429	0.952381
2 (random)	0.440000	1.007360	0.584320	0.440320	0.010826	0.010219	0.010144	0.943918
2 (selfing)	0.440000	1.022000	0.594145	0.445672	0.017716	0.017318	0.016945	0.977525
3 (random)	0.475222	1.017628	0.588245	0.436884	0.011477	0.010886	0.010667	0.948520
3 (selfing)	0.497065	1.044228	0.597535	0.447573	0.020509	0.020255	0.019397	0.987615

Table 3
Values of various quantities in Example 3 in the case of selfing with 1.1, 1.2 and 1.0 as fitness values and 0.2, 0.4 and 0.4 as initial frequencies of genotypes A_1A_1, A_1A_2 and A_2A_2, respectively.

Generation	p	$\bar{w}$	β_1	β_2	σ^2	σ_A^2	$\sigma_A^2/\bar{w}$	σ_A^2/σ^2
1	0.400000	1.100000	0.592857	0.521429	0.008000	0.002857	0.002597	0.357143
2	0.418182	1.074545	0.570490	0.513398	0.006261	0.002246	0.002229	0.393087
3	0.438240	1.062098	0.560606	0.507991	0.004790	0.002389	0.002249	0.498728

effects $\beta_1 = 0.55$, $\beta_2 = 0.45$ exactly fit these genotypic fitnesses whatever the genotype frequencies might be, these average effects do not change over time and are also not listed. The α version average effects do change over time and are listed. They can be calculated either from (1) or by subtracting the current value of $\bar{w}/2$ from the (fixed) β version values $\beta_1 = 0.55$ and $\beta_2 = 0.45$. The values of p, $\bar{w}$ and $\sigma_A^2/\bar{w}$ are also listed.

Example 2. In this example the fitnesses of the genotypes A_1A_1, A_1A_2 and A_2A_2 are assumed to be 1.2, 1.0 and 0.9, so that the fitnesses now exhibit dominance. Although dominance is a factor in the calculations, the focus in this example, as in Example 1, is on the effects of the mating scheme on the values of p, $\bar{w}$, σ_A^2 and the average effects in generations 2 and 3 as shown in Table 2.

Example 3. In this example the fitnesses of the genotypes A_1A_1, A_1A_2 and A_2A_2 are assumed to be 1.1, 1.2 and 1.0, so that the fitnesses now exhibit overdominance. The calculations involve only the case of selfing, since an analysis of the random-mating case is not illuminating. Values for this example are given in Table 3.

The values given in the three tables will be used later in the discussion about the two interpretations of the FTNS given in the next section.

4. Interpretations

What biological interpretation can be given to the partial change in mean fitness $\partial_{NS}\bar{w}$ as expressed on the right-hand side of (11)? Price initially described this to be the change in mean fitness "due to natural selection", as the suffix "NS" that he used clearly indicates, but he later expressed reservations about this interpretation. Similarly Fisher (1941, p. 57) used the words "ascribable to natural selection" in interpreting what Price writes as $\partial_{NS}\bar{w}$. A second interpretation is that it is the change "due to changes in gene frequencies". Fisher (1958, p. 37) uses the expression "rate of increase in [mean] fitness due to all changes in gene ratios" as an interpretation of $\partial_{NS}\bar{w}$. Thus he appears to use both this and the "due to natural selection" interpretations, perhaps interchangeably.

We reject these interpretations. Since natural selection affects the entire genome, it leads to changes not only in gene frequencies but also, as the examples in the previous section show, to significant changes in the average effect values caused by changes in the genotype frequencies. These also contribute to the change in mean fitness. Mean fitness changes also often depend on the mating scheme, as the examples show. When fitness depends on the genes at many loci, as is indeed the case in practice, natural selection also changes gametic frequencies and linkage disequilibrium values, and changes in mean fitness through natural selection derive from these changes also. Thus $\partial_{NS}\bar{w}$ is not the overall change in mean fitness due to natural selection: it is only one component of that change, and perhaps in some cases only a small proportion of that change.

This point is made clear from the values in Table 2 (for Example 2). Under both random mating and selfing the change in the frequency p of A_1 between generations 1 and 2 is the same (0.04). However, the change in mean fitness differs substantially in the two cases, being three times larger in the selfing case than in the random-mating case. This difference is reflected in the different values of β_1 and β_2 in the two mating schemes. Thus the difference in the change in mean fitness in these two cases can be explained entirely by the difference between the inter-generational changes in the numerical values of the average effects β_1 and β_2 and not at all by gene frequency changes. To consider a theory where average effect changes are assumed not to occur is to ignore a property of the evolutionary process.

Further, both in Examples 1 and 2, the generation 2 values of $\sigma_A^2/\bar{w}$ differ under the two mating schemes, being 0.010144 in the case of random mating and 0.016945 in the case of selfing in Example 2. Something other than a change in gene frequency has caused the change, so it cannot be claimed that this change is due only to changes in gene frequencies. It is due also to the different offspring generation values of β_1 and β_2 in the two mating schemes. Once again, to ignore the changes in β_1 and β_2, both with respect to changes in $\bar{w}$ and $\sigma_A^2/\bar{w}$, as is done in the above interpretations of the FTNS, is to ignore an important aspect of evolution. The statement of the FTNS is independent of the mating scheme. Thus to the extent that changes in gene frequencies depend on the mating scheme, the FTNS ignores an important aspect of the evolutionary process.

If one takes the view that changes in gene frequencies are primary and changes in average effects are secondary, the expression "due to changes in gene frequencies" can be justified. This is the view, for example, of Okasha (2008). We are unhappy about this interpretation also. The values of the average effects depend on the genotype frequencies (see Eq. (1)), so that the one-locus Eq. (17) should be written more fully as

$$\bar{w} = 2\{p\beta_1(P_{11}, P_{12}, P_{22}) + (1-p)\beta_2(P_{11}, P_{12}, P_{22})\}, \tag{18}$$

a form that explicitly recognizes that the average effects are functions of genotype frequencies, implying that changes in genotype frequencies imply changes in average effects. We are unhappy with the view expressed by Fisher and others that changes in average effects (as well as the mating scheme) can be thought of as part of the environment of the gene and that the partial change in mean fitness is then a change in a "constant environment" including a "constant genic environment". Price (1972) has similar reservations also: "… the device of treating non-additive gene effects as 'environment' [is disappointing]". We are even less happy with the view that the average effects can be viewed as being held constant by some unspecified extrinsic force. The above examples make it clear that changes in average effects are an important aspect of the evolutionary process.

As recalled in Lessard (1997), a constant genic environment is assured when "the change in genotypic frequencies [is] ascribable only to change in gene ratio" (Fisher, 1941, p. 56). In a context of multiple alleles at multiple loci, it has been shown (Lessard, 1997, Eq. (49)) that this occurs when the changes in genotypic frequencies are equal to their predicted values from the changes in gene frequencies, that is

$$\Delta P_g = \left[\Delta P_g\right]_\alpha = \frac{P_g \alpha_g}{\bar{w}}, \tag{19}$$

for every genotype g. This implies that

$$w_g = \bar{w} + \alpha_g = \beta_g, \tag{20}$$

for the fitness of every genotype g, which means no dominance at each locus and no epistasis between loci. This is the hidden connection between the two different expressions (11) and (15) for the partial change in mean fitness. In the former the additive genetic value of fitness is kept constant from the parental generation to the offspring generation, while in the latter this additive genetic value is assumed to capture all fitness differences. The above condition guarantees not only that the changes in genotype frequencies are linear with respect to gene frequencies, but also the change in mean fitness. It seems that what Fisher (1941, p. 57) meant by "a change ascribable to a change in gene frequency" is actually a change ascribable "only" to a change in gene frequency as if there would be no gene interactions at all. This is quite restrictive, though, and even misleading.

In Ewens (1989) the neutral expression "partial change" for $\partial_{NS}\bar{w}$ was introduced by analogy with a partial derivative to provide a term differing from the interpretations discussed above and not implying any specific evolutionary interpretation. This "neutral" expression does not, however, provide a meaningful biological interpretation. In the following two sections we turn to other properties of this change that do have a biological interpretation.

5. The mating scheme

Fisher (1941) criticized much of Sewall Wright's theoretical work because Wright almost always assumed random mating. In particular he showed that the well-known Wright equation

$$\Delta p = \frac{p(1-p)}{2\bar{w}} \frac{\partial \bar{w}}{\partial p} \tag{21}$$

for the change Δp in the frequency of some allele as a function of mean population fitness $\bar{w}$ holds only under random mating. Edwards (2000) has also criticized this equation from various further points of view. As noted above, random mating is not assumed in the statement of the FTNS.

A previous attempt (Ewens, 2010) has been given to derive a statement of the FTNS depending on the fact that it holds whatever the mating scheme, and which also involves $\partial_{NS}\bar{w}$. This is that "Nothing [in general] is known about the mating scheme, and therefore nothing is known [in general] about changes in mean fitness, …, since these … are determined in part by the mating scheme. This implies in turn that nothing can be stated in general about the long-term effects of selection as reflected by gene frequency changes. The FTNS isolates that part of the total change in mean fitness from one generation to the next about which something can be said independent of … the unknown mating scheme. This is that … component of the change in mean fitness that is independent of the mating scheme is $[\partial_{NS}\bar{w}]$".

Clearly this is not the interpretation of the FTNS that others have made, and we understand that it will probably not gain general acceptance. It can however be claimed that this is the only statement associated with the FTNS so far given that does not depend on the debatable assumptions discussed above, in particular that gene interactions and inter-generational changes in average effects may be ignored. The genotype frequencies depend on the mating scheme and so do the average effects, and consequently the mean fitness, unless alleles have frequency-independent additive effects on fitness, which is a very strong assumption.

6. The growth-rate theorem and the genic evolutionary potential

The FTNS with no dominance, called the growth-rate theorem and described as being elementary and no doubt known to Fisher by Edwards (1994), is the following. We imagine k "types" in a true-breeding population: the simplest genetical interpretation is k different allelic types at some locus in a haploid population, which is the application which we use. It is assumed throughout this paper, when discussing the haploid population version of the growth-rate theorem, that the fitness of any individual depends only on its allelic type at a single gene locus. The offspring generation frequencies of the alleles at this locus will in general differ from those in a parental generation, the changes being determined by the frequencies and fitnesses of the various alleles. We denote these allelic frequencies in some parental generation at conception by $p_1, p_2, \ldots, p_k$, with corresponding constant fitnesses $w_1, w_2, \ldots, w_k$, leading to a parental generation mean fitness of $\bar{w} = \sum_j p_j w_j$. The frequency of allele j in the offspring generation at conception will be the same as the frequency of this allele in the parental generation at reproduction, which is $p_j^* = p_j w_j/\bar{w}$. This increases the frequencies of the alleles that have the higher fitnesses and decreases the frequencies of the others. The offspring generation mean fitness is then $\bar{w}^* = \sum_j p_j^* w_j$ and the increase $\bar{w}^* - \bar{w}$ in mean fitness between the two generations is

$$\sum_j (p_j^* - p_j) w_j = \frac{\sigma^2}{\bar{w}}, \tag{22}$$

where

$$\sigma^2 = \sum_j p_j w_j^2 - \bar{w}^2 \tag{23}$$

is the variance of the allelic fitnesses in the parental generation. This is the growth-rate theorem. We call the left-hand side in Eq. (22) the "growth-rate evolutionary potential" and denote it by ρ_{GR}, so that

$$\rho_{GR} = \frac{\sigma^2}{\bar{w}}. \tag{24}$$

So long as σ^2 is positive the population has a positive evolutionary potential, that is a potential to increase in mean fitness through natural selection.

The naive statement of the Darwinian paradigm, that more fit parents leave more offspring who inherit in part the fitness of their parents, leading to a steady increase in population mean fitness, is similar to the statement of the growth-rate theorem. Fisher saw that the complexities of the diploid Mendelian hereditary system imply that a more subtle statement is needed. Finding such a statement was the aim of the FTNS. We note three properties of the growth-rate theorem as applied to the simple case of a haploid population where the fitness of any individual depends on the allele at some single given locus. First, one component of the change in mean fitness as given in Eq. (22) was found by multiplying each fitness w_j by a frequency at the time of reproduction. Second, the fitness w_j is equal to the average effect β_j of allele j as defined by an obvious haploid version of (4). Third, clearly a key component of the right-hand side in Eq. (22) is a variance.

We now develop the corresponding result for a diploid population. In doing so we generalize to the case where the fitness of any individual depends on its genes at all autosomal loci in its genome. (Slightly inaccurately we call this he "whole genome" generalization since the sex chromosomes are ignored.) We take the view of many authors, including Fisher, that the substance of evolution is a change in gene frequencies. However, as discussed below, this to us is a long-term viewpoint. Later we contrast this viewpoint with the statement of the FTNS.

Thus, in the diploid case, we seek to find a concept of evolutionary potential that relates to the various alleles in the genome and their frequencies. To do this we multiply the frequency of the whole-genome genotype g in the parental generation,which goes from P_g at the time of conception to

$$P_g^* = \frac{P_g w_g}{\bar{w}} \tag{25}$$

at the time of reproduction, not by its fitness w_g as in the "growth-rate" theorem, but instead by its "allele derived" fitness estimate

$$\beta_g = \sum_{i,j \text{ in } g} \beta_{ij}.$$

This estimate is obtained from the (β version) average effects of the alleles in genotype g, with β_{ij} occurring once, twice or not at all depending on how many times the allele A_j at locus i occurs in genotype g. Following the derivation of ρ_{GR} in (22), we then calculate

$$\sum_g P_g^* \beta_g - \sum_g P_g \beta_g = 2 \sum_i \sum_j (p_{ij}^* - p_{ij}) \beta_{ij} \tag{26}$$

as in the partial change (11), with p_{ij}^* being the frequency of allele A_j at locus i in the parental generation at the time of reproduction. Noting that this frequency is the same in the offspring generation at the time of conception under any system of mating (in the absence of sexual selection, segregation distortion and fertility differences) and proceeding as in Ewens (1989), we find that

$$\sum_g (\Delta^* P_g) \beta_g = 2 \sum_i \sum_j (\Delta^* p_{ij}) \beta_{ij} = \frac{\sigma_A^2}{\bar{w}}, \tag{27}$$

where $\Delta^* P_g$ and $\Delta^* p_{ij}$ are the changes in the frequencies of genotype g and allele j at locus i, respectively, from conception to reproduction in the parental generation. An alternative expression is

$$\sum_g \left[\Delta^* P_g\right]_\alpha w_g = \frac{\sigma_A^2}{\bar{w}}, \tag{28}$$

where

$$\left[\Delta^* P_g\right]_\alpha = \left[\Delta P_g\right]_\alpha = \frac{P_g \alpha_g}{\bar{w}}, \tag{29}$$

owing to Eqs. (13), (28), (33) and (50) in Lessard (1997), which means that the change in the frequency of genotype g as best predicted from the changes in gene frequencies is the same from conception to reproduction in the parental generation as from the parental generation to the offspring generation both at the time of conception. If the changes in genotype frequencies are equal to these predictions, then the change in mean fitness is equal to the change in a haploid population with constant fitnesses from one generation to the next which means no gene interaction.

From this we define the whole-genome diploid population "genic evolutionary potential" denoted by ρ, analogous to the "growth-rate" evolutionary potential ρ_{GR}, as the left-hand side member in Eq. (27), so that

$$\rho = \frac{\sigma_A^2}{\bar{w}}, \tag{30}$$

where σ_A^2 is the whole-genome additive genetic variance. Thus $\rho = \partial_{NS} \bar{w}$, hence providing another interpretation of $\partial_{NS} \bar{w}$. This is the whole-genome diploid population analogue of the "growth-rate" result (24). The word "genic" in genic evolutionary potential has the same meaning as Fisher's "genetic", that is "pertaining to genes". In the next section we list some of the properties of the genic evolutionary potential.

7. Properties of the genic evolutionary potential

Because of the relation of the genic evolutionary potential to σ_A^2 as given in Eq. (30), formulas involving ρ lead to formulas involving average effects. Perhaps the most convenient formula for calculating the additive genetic variance (see, e.g., Eq. (27) in Ewens, 1989) is

$$\sigma_A^2 = 2 \sum_i \sum_j p_{ij} \alpha_{ij} a_{ij}, \tag{31}$$

where a_{ij} is the average excess in fitness of allele j at locus i defined by

$$p_{ij} a_{ij} = \bar{w} \Delta^* p_{ij}, \tag{32}$$

with

$$\Delta^* p_{ij} = p_{ij}^* - p_{ij} = p'_{ij} - p_{ij} = \Delta p_{ij} \tag{33}$$

being the change in the frequency of that allele. This implies that

$$\rho = 2 \sum_i \sum_j (\Delta p_{ij}) \alpha_{ij}. \tag{34}$$

This expression brings out an interesting property of the genic evolutionary potential. The change $\Delta^* p_{ij}$ is not only the intra-generational change in gene frequency in the parental generation between the time of conception and the age of reproduction: it is also the inter-generational change in gene frequency between the parental generation and the offspring generation both at the time of conception. Thus it is independent of the mating scheme. This does not however imply that the mating scheme is irrelevant to the evolution of the population, as the examples in Section 3 show.

Suppose that the action of natural selection under a given mating scheme leads, over time, to a situation where σ_A^2 is zero, so that ρ is also zero. In this situation all average effects α_{ij} are 0, so that all average excesses a_{ij} are 0 (see, e.g., Eq. (26) in Ewens, 1989). Owing to (32), all changes in gene frequency $\Delta^* p_{ij}$ are 0. This implies that under that mating scheme there is no further potential for a change in gene frequencies. However, if the mating scheme then changes the genic evolutionary potential can become positive and gene frequencies can change. An example is given in the following paragraph.

Suppose in Example 3 (see Table 4) that the population mates at random and that the genotype frequencies have reached their

Table 4
Values of various quantities in Example 3 in the case of selfing with 1.1, 1.2 and 1.0 as fitness values and Hardy–Weinberg proportions 4/9, 4/9 and 1/9 as initial frequencies of genotypes A_1A_1, A_1A_2 and A_2A_2, respectively.

Generation	p	$\bar{w}$	β_1	β_2	σ^2	σ_A^2	$\sigma_A^2/\bar{w}$	σ_A^2/σ^2
1	0.666667	1.133333	0.566667	0.566667	0.004444	0.000000	0.000000	0.000000
2	0.666667	1.101961	0.561647	0.529647	0.004506	0.000723	0.000656	0.160415
3	0.676157	1.086833	0.556676	0.515731	0.003706	0.001301	0.001197	0.351005
4	0.690242	1.079633	0.553711	0.508856	0.003036	0.001609	0.001490	0.529836
5	0.706539	1.076550	0.552043	0.505126	0.002581	0.001757	0.001632	0.680703
6	0.723755	1.075661	0.551121	0.503012	0.002280	0.001810	0.001683	0.794183
7	0.741149	1.075948	0.550613	0.501784	0.002071	0.001806	0.001679	0.872071
8	0.758285	1.076851	0.550335	0.501062	0.001915	0.001766	0.001640	0.922291
9	0.774902	1.078060	0.550183	0.500634	0.001789	0.001705	0.001582	0.953408
10	0.790849	1.079402	0.550100	0.500380	0.001678	0.001631	0.001511	0.972268
20	0.907284	1.090729	0.550000	0.500002	0.000841	0.000841	0.000771	0.999850
50	0.994178	1.099418	0.550000	0.500000	0.000058	0.000058	0.000053	1.000000

random mating stable equilibrium values given by 4/9, 4/9 and 1/9 for the genotypes A_1A_1, A_1A_2 and A_2A_2, respectively. Then $\sigma_A^2 = 0$, and provided that random mating continues there is no genic evolutionary potential and there will be no change in gene frequencies in the following generations. If, however, the population were to start selfing, gene frequencies change and, provided that selfing continues, the frequency of A_1 steadily increases from 2/3 to 1. The genic evolutionary potential is positive throughout the process with values first increasing up to generation 6 and then decreasing afterward. It is interesting to note that, despite this, the mean fitness initially decreases from the generation 1 value 1.13333 as the frequency of the heterozygote A_1A_2 rapidly decreases, eventually reaching a minimum value of about 1.07566 (in generation 6), and then steadily increases as the number of heterozygotes becomes negligibly small and the more fit homozygote A_1A_1 gradually supplants the less fit homozygote A_2A_2, leading to a population mean fitness steadily increasing towards 1.1. The mean fitness values in generations 1–6 show that although a positive genic evolutionary potential implies a change in gene frequencies it does not necessarily imply an increase in mean fitness. It is also interesting that the mean fitness reaches a minimum in the same generation that the evolutionary potential reaches a maximum. It would be interesting to investigate whether this is a general phenomenon.

In the examples in Section 3 the fitness of any individual was assumed to depend on the genes at one locus only. This was sufficient to make the points described above. In the case where the fitness of an individual depends on many alleles at many loci in the genome, taking into account further complications due to epistasis and recombination would only reinforce those points. Selection and recombination have complicated effects on the population mean fitness when the fitness of any individual depends on entire genome, and only some of these effects are reflected in changes in gene frequencies.

Since $\sigma_A^2 \le \sigma^2$, where σ^2, defined by

$$\sigma^2 = \sum_g P_g\{w_g - \bar{w}\}^2, \tag{35}$$

is the total variance in fitness, the genic evolutionary potential must be less than or equal to $\sigma^2/\bar{w}$. Note that there would be equality in the case of true breeding owing to the growth-rate theorem. But genotypes are not transmitted, only genes are under several constraints from which the concept of genic evolutionary potential emerges. One can think of this as implying that complications deriving from the mating scheme and recombination properties of a diploid Mendelian population, all of which in part determine changes in gene frequencies as brought about by natural selection, imply that there is an inefficiency in increasing mean fitness as determined only by changes in gene frequencies, at least as compared to the simple haploid growth-rate theorem described above, where these complications do not arise. To quantify this deficit one can define a "genic efficiency ratio" by

$$\epsilon = \frac{\sigma_A^2}{\sigma^2}. \tag{36}$$

This ratio increases in our examples with selfing (see Tables 2–4), but decreases in Example 2 in the case of random mating (see Table 2). The difference $1 - \epsilon$ measures the proportional extent to which an increase in mean fitness as best predicted from changes in gene frequencies falls short of a maximum value because of the properties of the diploid Mendelian system referred to above.

The numerical value of the genic evolutionary potential ρ depends on the whole-genome genetic structures of the various individuals in the population. Differences in these structures almost always imply (population) linkage disequilibrium and departure from Hardy–Weinberg proportions. It is possible to calculate the marginal fitnesses of the genotypes at any gene locus and, through this, to calculate the marginal additive genetic variance at each locus. Despite implications to the contrary often make in the literature, the sum of these marginal additive genetic variances is hardly ever equal to the whole genome additive genetic variance as calculated above. The two will be equal if (and in practice essentially only if) all possible effects of dominance and epistasis, for all pairs of alleles at all possible pairs of loci, are zero. Given the likely architectured nature of many genomes, this is extremely unlikely to be the case. Nevertheless, if it were the case, the whole genome genic evolutionary potential would be the sum of the individual locus genic evolutionary potential values. This observation gives an insight into the relation between ρ and the architecture of the genome as determined by the possibly complex linkage disequilibrium and inbreeding properties.

8. Optimizing principles and the genic evolutionary potential

The FTNS is not a maximizing or an optimizing result, and was never stated as such by Fisher. An increase in some quantity does not imply a maximum possible increase (or rate of increase), nor does it imply that the effect of natural selection is to maximize population mean fitness at equilibrium points where gene frequencies do not change, and indeed epistatic and recombination effects ensure that mean fitness is almost always not maximized at equilibrium points.

Ewens (1992) gave an optimizing principle of natural selection associated with the concept of mean fitness. Since it is possible to change the population mean fitness by any chosen amount consistent with the current genotype fitness array by suitably arbitrary changes in gene frequencies, any optimizing principle must be a constrained one. In terms of the genic evolutionary potential concept this principle is as follows.

We define the diagonal matrix D to display the frequencies of all genes at all loci in the genome in some agreed ordering, the block-diagonal matrix X to display the corresponding frequencies of all

ordered one-locus genotypes at all loci in the genome in the same agreed ordering, and Q to be a matrix relating to two-locus frequencies of gene combinations, all calculated at the time of conception of the parental generation. (For more details about Q, see Castilloux and Lessard (1995).)

Let $\mathbf{d} = (d_{ij})$ be any vector of gene frequency changes (at all loci) consistent with the requirement that all gene frequencies at all loci and at all times be non-negative and, at each locus, add to 1. We call these "allowable" changes. In Ewens (1992) it was shown that a biologically natural (squared) distance measure between parental and offspring gene frequencies is $\mathbf{d}'\{2(D+X+Q)\}^{-1}\mathbf{d}$. (This distance measure differs by an unimportant factor of 2 from that given in Ewens (1992).) Then it can be shown that if $\boldsymbol{\delta} = (\Delta p_{ij})$ is the vector of changes in all the gene frequencies in the genome in the above agreed ordering that are brought about by natural selection, then

$$\boldsymbol{\delta}'\{2(D+X+Q)\}^{-1}\boldsymbol{\delta} = \frac{\sigma_A^2}{\bar{w}^2}, \tag{37}$$

where σ_A^2 is the whole-genome additive genetic variance, $\bar{w}$ is the (whole-genome) population mean fitness and both σ_A^2 and $\bar{w}$ are calculated at the time of conception of the parental generation. In terms of the genic evolutionary potential concept and formula (34) this leads to the following statement.

Theorem. *Of all allowable changes* $\mathbf{d}$ *in gene frequencies that lead to the same distance between parental and offspring gene frequencies, that is that satisfy*

$$\mathbf{d}'\{2(D+X+Q)\}^{-1}\mathbf{d} = \sigma_A^2/\bar{w}^2, \tag{38}$$

the natural selection vector $\boldsymbol{\delta} = (\Delta p_{ij})$ *maximizes*

$$2\sum_i\sum_j d_{ij}\alpha_{ij}, \tag{39}$$

so that the genic evolutionary potential (equivalently, Price's $\partial_{NS}\bar{w}$*) is the maximum value.*

9. Long-term gene frequency changes

We agree with those who view long-term changes in gene frequencies as the substance of evolution. It has been noted above (and the FTNS relies on this fact) that changes in gene frequencies between a parental generation at its time of reproduction and the resultant offspring generation at its time of conception do not depend on the mating scheme adopted by the parental generation. However, long-term changes in the frequencies of genes in the presence of dominance and epistasis, indeed changes over two or more generations, do depend on the mating scheme sometimes significantly, as the examples in Section 3 show. In more realistic examples they also depend on the recombination structure in the genome and the linkage disequilibrium structure of the population. These dependencies are not captured, nor are they intended to be captured, by the FTNS, which is then not a long-term result.

10. Summary

With the current emphasis in population genetics on data analysis and the retrospective theory, it might seem old-fashioned to continue to discuss the FTNS. There are two reasons for discussing it in this paper. First, to the extent that Fisher regarded the FTNS as the centerpiece of his work in evolutionary genetics and its still-continuing misinterpretation in parts of the literature, continuing clarifications are needed. Second, with the resurgence of a discussion of the FTNS in the context of inclusive fitness, the theorem remains of interest.

Our major two problems with the FTNS are as follows. First, as discussed in the previous section, the FTNS says nothing about long-term gene frequency changes, since it ignores too much relevant information to be able to do this. It cannot be used to calculate gene frequency changes even two generations in advance. To the extent that long-term changes in gene frequencies are the substance of evolution it provides little long-term evolutionary information. On the other hand, the theorem does not claim to calculate long-term gene frequency changes, so that this comment is not a criticism of the theorem. Second, even as a "one generation in advance" result, the unrealistic assumptions made in arriving at and even interpreting the theorem diminish its value.

The incorrect interpretation of the FTNS, that mean fitness increases as time goes on (or at least remains unchanged) can still be found in some textbooks and research papers. If however that result really were true, then it would provide a long-term relevance for the FTNS. It can be argued that analyses of the total change in mean fitness and its approximation under weak selection (Kempthorne, 1957; Kimura, 1958; Nagylaki, 1976, 1993; Charlesworth, 1994), determining the circumstances in which population mean fitness can be expected to increase, will eventually prove to be a more fruitful approach to understanding evolution as a genetic process than an approach following the lines of the FTNS.

Despite these comments, the importance of the FTNS in introducing the relevance of the additive genetic variance in fitness for evolutionary considerations is undeniable. The relevance of the additive genetic variance in plant and animal breeding programs via the concept of heritability is central. However, this can be regarded as the exception that proves the rule concerning the thesis of this paper. Strictly speaking, heritability calculations are for one generation in advance only. They do not determine the properties of evolution over hundreds of generations, and in this they share the comments made above about the FTNS.

Further, the central observation which leads to the FTNS, namely that a naive statement of the Darwinian theory is not appropriate for diploid populations and that a more sophisticated statement is needed, is of first-rate importance. Whatever one's views about the FTNS might be, this observation by Fisher is arguably the most important revision of the Darwinian paradigm since Darwin's time.

Acknowledgments

We acknowledge the comments of various referees which have improved the first version of this paper. This research was supported in part by NSERC of Canada, Grant No. 8833.

References

Castilloux, A.M., Lessard, S., 1995. The fundamental theorem of natural selection in Ewens' sense (Case of many loci). Theor. Popul. Biol. 48, 306–315.
Charlesworth, B., 1994. Evolution in Age-Structured Populations, second ed. In: Cambridge Studies in Mathematical Biology, vol. 1. Cambridge University Press, Cambridge.
Crow, J.F., 1979. Gene frequency and fitness change in an age-structured population. Ann. Hum. Genet. 42, 355–370.
Edwards, A.W.F., 1994. The fundamental theorem of natural selection. Biol. Rev. 69, 443–474.
Edwards, A.W.F., 2000. Sewall Wright's equation $\Delta p = \{p(1-p)\partial w/\partial p\}/2w$. Theor. Popul. Biol. 57, 67–70.
Ewens, W.J., 1989. An interpretation and proof of the fundamental theorem of natural selection. Theor. Popul. Biol. 36, 167–190.
Ewens, W.J., 1992. An optimizing principle of natural selection in evolutionary population genetics. Theor. Popul. Biol. 42, 333–346.
Ewens, W.J., 2010. What is the gene trying to do? British J. Philos. Sci. 62, 155–176.
Fisher, R.A., 1941. Average excess and average effect of a gene substitution. Ann. Eugen. 11, 53–63.
Fisher, R.A., 1958. The Genetical Theory of Natural Selection, second revised edition. Dover, New York.
Grafen, A., 2015. Biological fitness and the fundamental theorem of natural selection. Am. Nat. 186, 1–14.

Kempthorne, O., 1957. An Introduction to Genetic Statistics. Wiley, New York.
Kimura, M., 1958. On the change of population fitness by natural selection. Heredity 12, 145–167.
Lessard, S., 1997. Fisher's fundamental theorem of natural selection revisited. Theor. Popul. Biol. 52, 119–136.
Nagylaki, T., 1976. The evolution of one- and two-locus systems. Genetics 83, 581–600.
Nagylaki, T., 1993. The evolution of multilocus systems under weak selection. Genetics 134, 627–647.
Okasha, S., 2008. Fisher's fundamental theorem of natural selection—a philosophical analysis. British J. Philos. Sci. 59, 319–351.
Price, G.R., 1970. Selection and covariation. Nature 227, 520–521.
Price, G.R., 1972. Fisher's fundamental theorem made clear. Ann. Hum. Genet. 36, 129–140.