Journal of Theoretical Biology 472 (2019) 77–83

Contents lists available at ScienceDirect

Journal of Theoretical Biology

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

The left-hand side of the Fundamental Theorem of Natural Selection: A reply

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

[a] Département de mathématiques et de statistique, Université de Montréal, Montréal, Québec H3C 3J7, Canada
[b] Department of Biology, The University of Pennsylvania, Philadelphia, PA 19104, USA

ARTICLE INFO

Article history:
Received 4 December 2018
Revised 12 March 2019
Accepted 12 April 2019
Available online 13 April 2019

Keywords:
Additive genetic variance
Average effects
Constant genic environment
Partial change in mean fitness

ABSTRACT

In a recent paper, Grafen (2018) discussed the left-hand side in the equation stating Fisher's (1930, 1958) "Fundamental Theorem of Natural Selection" (FTNS). Fisher's original statement of the FTNS is, in effect, "The rate of increase in fitness of any organism is equal to its genetic variance in fitness at that time" with the rate of increase in fitness understood as the one "due to all changes in gene ratios" (Fisher, 1930, p. 35). For purposes of exposition, Grafen (2018) considered what is today called the analogous discrete-time model, and restated the FTNS on p. 181 as "The increase in population [mean fitness] due to changes in gene frequencies [is equal to the] additive genetic variance in fitness [divided by the] mean fitness". Allowing for the fact that Grafen's statement of the FTNS relates to a discrete-time model, his statement is in effect a discrete-time version of Fisher's. It has however been widely accepted for many years, ever since Price's (1972) deep analysis of the FTNS, that Fisher's wording does not correctly describe the content of the FTNS. The same is therefore true of Grafen's statement. The confusion caused by these misstatements is unfortunate and adds to a continuing misunderstanding of the FTNS, whose source can also be found in Fisher's (1941) own explanation. Our purpose is to review the detailed analysis of the calculations leading to the FTNS to clarify the points at issue.

1. Introduction

The Fundamental Theorem of Natural Selection (FTNS) (Fisher, 1930; 1958) was an attempt to capture the effect of natural selection in Mendelian populations. A direct consequence of selection is to change gene frequencies, but with many different possible combinations of genes and complex mating schemes in diploid populations, the effect of selection on gene frequencies becomes rapidly intricate. On the other hand, it is of interest to describe the effect of changes in gene frequencies on the population state and its evolution over time. This is the aim of the FTNS. But a diploid population evolves under multiple effects of multiple types, genetic and environmental. Nevertheless, it is imagined that the effect of changes in gene frequencies can be isolated from all other effects. All these are combined together in what is called the effect of the environment of the genes. When this effect on the mean fitness is removed, what is left is generally accepted to give the FTNS. Therefore, the FTNS concerns a partial change in mean fitness. It took some time for this to be understood (Price, 1972; Ewens, 1989; Lessard, 1997), but it is now widely accepted. Apart perhaps the definition of fitness in age-structured populations and more generally in class-structured populations (Grafen, 2015a; Lessard and Soares, 2016), what remains a matter of debate over the FTNS is the interpretation and biological significance of the partial change in mean fitness. It is widely viewed as the change due to natural selection, and more precisely due to changes in gene frequencies by natural selection. In a recent paper (Ewens and Lessard, 2015), we have argued that this partial change does not actually take into account all effects of changes in gene frequencies. As a result, it may be of limited interest to describe the effect of natural selection on long-term as well as short-term evolution. More recently, Grafen (2018) has challenged this point of view, revisiting the partition of the mean fitness and claiming that the FTNS captures the effect of changes in gene frequencies on mean fitness and remains fully meaningful to biologists. We will review the arguments and see what are exactly the issues.

2. Fisher's FTNS

We initially discuss the case focused on by Grafen (2018) for which the fitness of any individual depends on the genes this individual has at one locus only, say locus "*A*". We call this the "one-locus" case, and generalize to the "whole-genome" case with

* Corresponding author.
E-mail address: lessards@dms.umontreal.ca (S. Lessard).

https://doi.org/10.1016/j.jtbi.2019.04.014
0022-5193/

multiple loci later, as is necessary since Fisher's FTNS is a whole-genome result. Suppose that the possible alleles at the "A" locus are $A_1, A_2, \ldots, A_k$ and that the viability fitness of an A_iA_j individual is w_{ij}. Suppose also that at the beginning of a parental generation, in a framework of an infinite population undergoing discrete, non-overlapping generations, the frequency of the (ordered) genotype A_iA_j is P_{ij}. Then the mean fitness $\bar{w}$ of the population at that time is $\sum_i \sum_j P_{ij} w_{ij}$.

Fisher's analysis focused on the concept of the average effect of an allele (or a gene in Fisher's terminology). At the beginning of the parental generation, the average effects of $A_1, A_2, \ldots, A_k$ are defined as those values of $\beta_1, \beta_2, \ldots, \beta_k$ which minimize the expression $\sum_i \sum_j P_{ij}(w_{ij} - \beta_i - \beta_j)^2$. In other words, the average effects are defined so that $\beta_i + \beta_j$, called the "additive genetic value", provides the best linear fit (in the sense of weighted least squares) to the fitness w_{ij} of A_iA_j as determined by the alleles A_i and A_j in that genotype. Since these average effects are characteristic of the alleles $A_1, A_2, \ldots, A_k$, they were viewed by Fisher as being central to his quantification of evolution by natural selection. They can indeed be thought of as the "fitnesses" of the alleles $A_1, A_2, \ldots, A_k$, a view that Fisher held, as do we. Regression theory shows that the average effects also have the property that the mean fitness $\bar{w}$ is identical to $\sum_i \sum_j P_{ij}(\beta_i + \beta_j)$. In other words, the mean fitness $\bar{w}$ of the population can be found from the "fitnesses" of the alleles in it.

The total variance σ^2 in fitness at the beginning of the parental generation is given by $\sum_i \sum_j P_{ij}(w_{ij} - \bar{w})^2$. That part of this variance which is explained by the average effects $\beta_1, \beta_2, \ldots, \beta_k$ is called the "additive genetic variance" (for Fisher, the "genetic variance"), and is denoted by σ_A^2. If it so happens that $w_{ij} = \beta_i + \beta_j$ for all (i, j) combinations, then the average effects explain the genotypic fitnesses exactly and $\sigma_A^2 = \sigma^2$. If $w_{ij} \neq \beta_i + \beta_j$ for one or more (i, j) combinations, then the average effects do not explain the genotypic fitnesses exactly and $\sigma_A^2 < \sigma^2$. In regression terms, σ_A^2 is a regression sum of squares.

It is a central and crucial observation that the average effects are not fixed constants, but depend on the set of genotype frequency values $\{P_{ij}\}$, and thus in general will differ from one set of genotype frequency values to another, in particular from a parental generation to an offspring generation. The only exception to this arises in the case discussed above where $w_{ij} = \beta_i + \beta_j$ for all (i, j) combinations. In this case (and in this case only), $\beta_1, \beta_2, \ldots, \beta_k$ are fixed constants in that they do not change from one generation to the next.

Define ΔP_{ij} as the change in the frequency of the genotype A_iA_j from the beginning of the parental generation to the beginning of the offspring generation, and correspondingly Δp_i as the change in the frequency of the allele A_i during the same time period. Then it was shown by Ewens (1989) and many subsequent authors that

$$2\sum_i (\Delta p_i)\beta_i = \sigma_A^2/\bar{w}. \tag{1}$$

We present this equation first since it relates to changes in gene frequencies, which were of central interest to Fisher. On the way to developing (1), it was shown that

$$\sum_i \sum_j \left(\Delta P_{ij}\right)(\beta_i + \beta_j) = \sigma_A^2/\bar{w}. \tag{2}$$

Note that no particular mating scheme was assumed in deriving these equations. In particular, they do not rely on the usual assumption of random mating.

Fisher's statement of the FTNS implied a continuous-time model. The continuous-time analogue of Eq. (1) is

$$2\sum_i \left(\frac{dp_i}{dt}\right)\beta_i = \sigma_A^2, \tag{3}$$

while

$$\sum_i \sum_j \left(\frac{dP_{ij}}{dt}\right)(\beta_i + \beta_j) = \sigma_A^2 \tag{4}$$

is the continuous-time analogue of Eq. (2).

Fisher's FTNS is a whole-genome result, so that Eqs. (1)–(4) are not sufficient to discuss the FTNS in full detail. We therefore state the whole-genome generalization of the above equations. We list the large number of genotypes at multiple loci in some agreed order as $G_1, \ldots, G_g, \ldots, G_m$. We denote the fitnesses of these genotypes by $w_1, \ldots, w_g, \ldots, w_m$ respectively, and assume that at the beginning of the parental generation they have respective frequencies $P_1, \ldots, P_g, \ldots, P_m$. Then the population mean fitness at that time is $\bar{w} = \sum_g P_g w_g$. The average effects of the various alleles at all loci in the genome at that time are found by minimizing $\sum_g P_g(w_g - \beta_g)^2$, where β_g is defined as the sum of the average effects of all alleles at all loci in the genome in genotype g, with the average effect of any allele at any locus being counted once, twice or not at all in β_g depending on whether that allele occurs once, twice or not at all in the whole-genome genotype g. Thus β_g is the whole-genome generalization of the additive genetic value $\beta_i + \beta_j$ described above for the one-locus case. The whole-genome additive genetic variance, for which we again use the notation σ_A^2, now of course having a whole-genome interpretation, is that part of the total variance in fitness $\sum_g P_g(w_g - \bar{w})^2$ which is explained by the average effects of all the alleles, at all loci, in the genome. It is therefore the whole-genome generalization of the additive genetic variance discussed above for the case where the fitness of any individual depends on the genes this individual has at one locus only. As in the one-locus case, σ_A^2 is a regression sum of squares.

Define ΔP_g as the change in the frequency of the whole-genome genotype g from the beginning of the parental generation to the beginning of the offspring generation, and correspondingly Δp_{ku} as the change in the frequency of the allele A_{ku} at gene locus u during the same time period. We denote the average effect of the allele A_{ku}, derived from the minimization procedure described above, by β_{ku}. Then in the discrete-time model, the whole-genome generalization of (1) is

$$2\sum_u \sum_k (\Delta p_{ku})\beta_{ku} = \sigma_A^2/\bar{w}. \tag{5}$$

Similarly the statement

$$\sum_g (\Delta P_g)\beta_g = \sigma_A^2/\bar{w} \tag{6}$$

is the generalization of (2).

The continuous-time model versions of (5) and (6), found after an analysis parallel to that which leads to (5) and (6) in a discrete-time model, are

$$2\sum_u \sum_k \left(\frac{dp_{ku}}{dt}\right)\beta_{ku} = \sigma_A^2 \tag{7}$$

and

$$\sum_g \left(\frac{dP_g}{dt}\right)\beta_g = \sigma_A^2, \tag{8}$$

respectively. Note that, again, random mating is not assumed in deriving these equations. By analogy with the concept of a partial derivative, the left-hand sides in (1), (2), (5) and (6) have been called "partial" changes in mean fitness, and the left-hand sides in (3), (4), (7) and (8) have been called "partial" rates of change in mean fitness. Eqs. (1)–(8) have been known for 30 years and are accepted statements of the FTNS, respectively, in the one-locus/whole-genome and discrete-time/continuous-time cases.

What do these equations imply? We first discuss Fisher's claim for the continuous-time model that the left-hand side in (7) is the

"rate of increase in fitness due to all changes in gene [frequencies]" (Fisher, 1930, p. 35, Fisher, 1958, p. 37). We have replaced Fisher's "ratio" by "frequencies" in the statement since we allow for an arbitrary number of alleles at each locus. At first sight this statement seems plausible. The left-hand side in (7) does indeed explicitly contain the set $\{dp_{ku}/dt\}$ of all rates of change in the frequencies of all alleles at all loci in the genome. The difficulty with Fisher's statement, repeated several times by several authors since (Price, 1972), is that in practice the average effects $\{\beta_{ku}\}$ appearing in (7) are almost certainly not constant over time. This is because their values depend on gene frequencies so that they will themselves change as gene frequencies change. Thus the left-hand side in (7) is not the rate of change in mean fitness due to changes in gene frequencies, since it ignores changes in the average effects $\{\beta_{ku}\}$. These changes might indeed often be small, but one of the purposes of the calculations in Ewens and Lessard (2015) was to show that they need not be small. In fact, it has been known for more than 50 years (Moran, 1964) that, when fitness depends on the alleles at just two loci, changes in the average effects can be so large relative to changes in gene frequencies that the population mean fitness can decrease over time. Thus, if taken literally, Fisher's wording does not describe exactly the content of (7).

3. Grafen's analysis

The left-hand side in the statement of the FTNS is the subject of Grafen's (2018) discrete-time analysis paper, as the title of that paper indicates. Before discussing this paper we make a general comment on it. We are disappointed in Grafen's (2018, p. 175–176) attacks on "mathematical population geneticists", who are accused of "wrongly proving [the FTNS] false", "wrongly proving it requires more assumptions than Fisher admitted", "doubting its biological significance", distracting biologists by "serious misreading of the theorem", that there is a "lack of appreciation by mathematical population geneticists (e.g., Ewens and Lessard, 2015)", and so on. All these statements are unfortunate since what should matter is scientific truth irrespective of the background of the researchers. It is certainly true that various mathematicians (and also many biologists) have misunderstood the statement of the FTNS. In our opinion, Grafen's paper adds to these misunderstandings. Moreover, insofar as Grafen's comments relate to Ewens and Lessard (2015), they are unwarranted and incorrect. The true situation is provided below.

For purposes of exposition, Grafen (2018) only analysed the case where the fitness of any individual depends on his genetic make-up at a single locus, with only two alleles possible at this locus. For the whole-genome analysis, he relies on, and quotes, the results given in Ewens and Lessard (2015), specifically (6) above. Again for purposes of exposition, he only considered a discrete-time model. Grafen's analysis proceeds via regression coefficients, but we proceeded above via average effects, since operating in this way makes the analysis more transparent and closer to Fisher's work. (The average effects are, however, identical to Grafen's regression coefficients, as both he and we note, so this is only a matter of style.)

Grafen (2018) subdivides the total change in mean fitness from one year to the next, in effect from one generation to the next, into three components. He calls these the "additive genetic component", the "non-additive genetic component" and the "phenotype-genotype component", and denotes them Δ_{AG}, Δ_{NAG} and Δ_{PGL}, respectively. He states that "So far as I know, this particular partition [of the total change in mean fitness as $\Delta_{AG} + \Delta_{NAG} + \Delta_{PGL}$] has not been proposed before". The phenotype-genotype component is important indeed. It relates to the case where fitness values change from one generation to the next. Certainly such changes are very likely to occur, but then the fitness values in the offspring generation will usually not be known in advance. Changes in fitness values are associated in the classical literature to changes in the environment excluding the genic environment (Fisher, 1930, 1941, see, e.g., Crow and Kimura, 1970, Eqs. (5.6.15) and (5.6.19)). In all analyses of the FTNS, including Fisher's, attention focuses on the case of constant fitness values. In this case, Grafen's (2018) partition reduces to two components. The formula for the additive genetic component Δ_{AG} is, in Grafen's (2018) notation, $\sum_g \left(\Delta p_g\right)\hat{x}_g$, which is $\sum_i \sum_j \left(\Delta P_{ij}\right)(\beta_i + \beta_j)$ in the present notation, that is the left-hand side in (2) as in Ewens (1989). As for the non-additive genetic component Δ_{NAG}, it is what is left in the case of constant fitness values once the additive genetic component is removed. As the title of Grafen's paper, and of ours, indicates, the focus of his paper, and of ours, is on these two components, so only these are discussed here.

Grafen (2018) states the one-locus FTNS via his Eq. (2). He denotes the left-hand side in this equation by $\Delta_{NS}\bar{w}$ (using a subscript found in Price, 1972, and Ewens and Lessard, 2015), but this is simply a new notation for his $\Delta_{AG}\bar{w}$, which as noted above is, in our present notation, $\sum_i \sum_j \left(\Delta P_{ij}\right)(\beta_i + \beta_j)$. The right-hand side in his Eq. (2) is $\sigma_A^2/\bar{w}$. Therefore, his Eq. (2) is identical to our Eq. (2) above. Thus, his Δ_{AG} is our $\sum_i \sum_j \left(\Delta P_{ij}\right)(\beta_i + \beta_j)$. Thus Grafen's partition of the total change in mean fitness in discrete time with constant fitness values as $\Delta_{AG} + \Delta_{NAG}$ is not really new, except perhaps in the wording, the notation or the graphical representation. It is the same partition as was made in Ewens (1989) (see, e.g., Lessard, 1997, Sections 5 and 8.2 and references therein, for similar discrete-time as well as continuous-time versions).

The fact that Grafen's Eq. (2) agrees with the standard statement of the FTNS in discrete time in the one-locus case as given by (1) and (2) above at least implies that there is no disagreement on what that statement is. However, Grafen (2018, p. 181) then rewrites his Eq. (2) as "The increase in population $\bar{w}$ due to changes in gene frequencies [is equal to] the additive genetic variance in fitness [divided by the] mean fitness". This is justified only by the facts that it is the change predicted by a linear regression on gene frequencies and that the corresponding change for any quantitative trait in the case of no fitness differences is zero owing to the Hardy–Weinberg law (Grafen, 2018, p. 177–178). However, Grafen never presents a formula for the change in mean fitness due to changes in gene frequencies alone. His verbal presentation of the FTNS (and its explanation) is the discrete-time analogue of Fisher's, and is similarly incorrect as stated, since changes in gene frequencies cause changes in average effects, not to mention changes in the genic environment. It has been suggested in the literature that the FTNS deals only with the "direct effect" of changes in gene frequencies by natural selection (see, e.g., Frank and Slatkin, 1992; Lee and Chow, 2013; Okasha, 2018), or with "changes in genotypic frequencies directly consequent on changes in gene frequencies" (Lessard, 1997), under the view that all other changes including changes in average effects, and more generally in the environment of the genes, are indirect. This verbal explanation is certainly better than Fisher's and Grafen's statements. Moreover, explicit algebraic expressions are important to understand the exact meaning of the FTNS and its interpretation.

Unfortunately, Grafen (2015a) appears to be quite explicit in making the same incorrect interpretation of the FTNS in an age-structured population as that discussed above, namely that the partial change in mean fitness obtained by keeping fixed the average effects is the change due to changes in gene frequencies. Here, he considers a "breeding value" ϕ_g for any genotype g, which, for that genotype, is (in our notation) exactly β_g in (6) minus the mean. He then goes on to say that this sum is to be regarded as "fixed over time" (Grafen, 2015a, p. 5). This is precisely the mis-

understanding referred to above: this sum *does* change over time as gene frequencies change. To the extent that Grafen's (2015a) approach depends on this assumption, it will arrive at the "age-structure" generalization of (7), or equivalently of (8), with an incorrect interpretation. It will then be subject to the same criticisms of Fisher's and Grafen's previous statements as were made above, that the expression derived will not be the change in mean fitness "due to natural selection" understood as "due to all changes in gene frequencies" as claimed (Grafen, 2015a, p. 7).

4. What are the implications?

Once the mathematical forms of the various FTNS equations have been established it is not only appropriate, but also necessary, for biologists to assess what the biological usefulness of these equations might be. But it is not unreasonable for mathematical population geneticists, having established the equations, to comment on them also. Price (1972) was the first to do so. Since then, mathematical population geneticists have pointed out several times that these equations do not agree with Fisher's (1930) verbal statement of the theorem. Nor would they agree with Grafen's (2018) statement of the theorem. There is no misreading of the theorem by mathematical population geneticists. Grafen's comments and his verbal statement of the FTNS given in his Eq. (2) add nothing to the ongoing discussion, and indeed only obfuscate the issue.

Mathematicians have gone further, in an attempt at being constructive, to find a biologically useful interpretation of the left-hand sides in Eqs. (5)–(8). For example, the first sentence in the "Interpretations" section of Ewens and Lessard (2015, p. 62) is: "What biological interpretation can be given to [this left-hand side]?" We show, for example, that this left-hand side is not affected by changes in the mating scheme in discrete time since this scheme does not change the gene frequencies, a point of interest to Fisher. Mathematicians have derived the whole-genome version of the FTNS and confirmed Fisher's insight that it is a whole-genome statement. Nagylaki (1976) initiated a line of research assessing the circumstances under which changes in mean fitness due to changes in average effects are small, so that Fisher's claim is "almost" correct. On the other hand, some authoritative statements made by biologists have contributed to a negative view of the FTNS. In their definitive exposition of the theory of evolution of quantitative traits, Walsh and Lynch (2018) state in effect that: "The bold, sweeping classical interpretation of Fisher's statement is replaced by an exact [...] result that is absolutely correct, but not really useful". This is a stronger negative statement than mathematicians have made about the FTNS, in contrast to Grafen's assertions.

5. Final thoughts

Various simplifying assumptions were made by Fisher (1930) in developing the FTNS. For example, random changes in gene frequencies (random drift) and the effect of the sex chromosomes are ignored. This is perfectly reasonable since no useful analysis could be made when all possible complicating factors are taken into account. His work has never been criticized for this and indeed most analyses of the FTNS, by most authors, make the same simplifying assumptions. On the other hand, non-random mating, whole-genome fitness determination and age effects with non-overlapping generations, which were explicitly not excluded by Fisher, have received much attention by mathematical population geneticists (see, e.g., Kempthorne, 1957; Kimura, 1958; Ewens, 1989; Lessard, 1997). Thus Grafen's (2018, p. 175) claim that they have been "wrongly proving that [the FTNS] requires more assumptions than Fisher admitted" is untrue.

Apart from his comments referred to above, Grafen (2018, p. 176) claims that the main interest of Ewens and Lessard (2015) "seems to be in total change in the mean fitness". This is not the truth. This is shown by the very fact that the focus of our calculations is precisely the same as Grafen's, namely the additive genetic variance and its relation to that part of the total change in mean fitness that has some evolutionary importance. More precisely, the quantities that Grafen and we focus on (his additive genetic component Δ_{AG} and the left-hand side in our Eq. (5)) are mathematically identical, differing only in notation, and neither is the total change in mean fitness nor the partial change due to changes in gene frequencies. This is that part of the total change in mean fitness for which we and others are attempting to find some biological relevance. Grafen's various comments on Ewens and Lessard (2015), and on mathematical population geneticists generally, are inaccurate.

Next, Fisher (1941) referred to the "environment" in which a gene finds itself, and discussed for example the concept of a "constant genic environment". He considered the rest of the genome besides the locus at which any gene arises, and also the population mating scheme (which might change from one generation to another), as part of the environment in which the gene finds itself, in addition to the physical and ecological parts that may be responsible for stochastic as well as frequency-and-density-dependent variations (see Appendix A). It is now widely recognised that Fisher's (1930) FTNS concerns the change in mean fitness obtained by keeping fixed the environment, including the genic environment. This change corresponds to a partial change in mean fitness and does not require restrictive assumptions such as weak selection and random mating. Nevertheless, the conditions for a constant genic environment happen to correspond exactly to conditions that prevail under these assumptions (see Appendix B). Then all statistical interactions between genes can be neglected (at least in the long run) and the change in a mean measurement can be predicted from average effects of genes and changes in their frequencies. The biological significance of the FTNS when these assumptions do not closely hold is unclear, however, since all effects depend on the population state, and in particular on gene frequencies. The partial change it concerns may be far from being the dominant component of the total change in mean fitness by selection in natural populations. It is a statistical simplification that tries to explain as much change as possible from the perspective of individual genes. It can become an oversimplification, and even a misrepresentation, of a complex biological reality in presence of interactions.

Moreover, and perhaps most important, there seems to be a circularity in Fisher's (1941) interpretation of the FTNS. The original argument has been presented in a continuous-time setting with two alleles at one locus (see Appendix A for details). Put in a discrete-time setting with multiple alleles at a single locus, the argument goes as follows.

First, the increase in mean fitness from one generation to the next is said to be ascribable only to changes in gene frequencies (or ascribable to natural selection) when it is given by the additive genetic variance in fitness divided by the mean fitness. The latter condition holds when changes in genotype frequencies satisfy Eq. (22). But then this condition involving Eq. (22) is used as the definition of a constant genic environment which, finally, is seen as justifying the initial attribution of increase in mean fitness to changes in gene frequencies alone. This is the reasoning in Fisher (1941) to conclude that the additive genetic variance in fitness divided by the mean fitness is the increase in mean fitness from one generation to the next ascribable only to changes in gene frequencies. This looks like a circular argument.

Actually, Fisher's (1941) explanation of the FTNS relies on a property of the average effects of any measurement, here fitness (call it measurement A) that is explained by the average effects

of another measurement, namely the relative change in genotype frequency (call it measurement B, see Appendix A). A constant genic environment corresponds to a measurement B that is additive, which is equivalent to saying a measurement that is determined by its average effects. This ensures that the change in the mean measurement of A is such that the non-additive genetic component Δ_{NAG}, in Grafen's (2018) terminology and notation, vanishes even if the measurement A is not additive. Then, necessarily, the change in the mean measurement of A, assuming that it is frequency-independent, is equal to the additive genetic component. However, this does not mean that the average effects of A nor the residual addends (dominance and epistasis effects) are kept constant. It means only that the change in the mean measurement of A in a constant genic environment is like the change in a population of genes with the same average effects, which are kept constant, and without any population structure. This corresponds to the growth-rate theorem for a haploid population (Edwards, 1994) with fitnesses given by the average effects.

Moreover, the partial change in mean fitness in the FTNS is not generally the change due to changes in gene frequencies, since these modify not only the average effects but also the residual addends. It is actually the change due to changes in the relative genotype frequencies as best predicted from the gene frequencies in each genotype (Lessard, 1997). In a partial selfing model in the limit of weak selection, for instance, the changes in genotype frequencies are functions of the changes in gene frequencies only and the change in mean fitness due to changes in gene frequencies is the total change, not just its additive genetic component (see Appendix B). Let us recall that the failure of the FTNS to capture satisfactorily the effect of natural selection in populations with inbreeding was a motivation for Hamilton (1964) to introduce the concept of inclusive fitness.

Finally, Grafen (2018) takes number of offspring as the measure of fitness in a context of non-overlapping generations but makes the surprising claim that "it is the accompanying definition of fitness that should be admired and recognised as the hallmark of genius in the fundamental theorem" (p. 187) in a context of age-structured populations with reference to Grafen (2015a). This definition is based on birth and death rates that depend on genotype and age but with reproductive values that depend only on age, and is such that the mean fitness at every age is the same in agreement with a neutrality principle recently stated for class-structured populations (Grafen, 2015b). It takes some imagination to find all this in Fisher's work. The partial change in mean fitness was obtained with the more classical Malthusian parameter as measure of genotypic fitness and reproductive values that depend on genotype and age in Lessard (1997). It was shown in Lessard and Soares (2016, 2018) that the two definitions of fitness agree in the limit of weak selection with a time scale shorter for changes in age distribution than for changes in gene frequencies.

Acknowledgments

This research was supported in part by NSERC of Canada, Grant no. 8833 to S. Lessard. We also appreciate the constructive comments of two anonymous referees.

Appendix A. Constant genic environment in the FTNS

The key argument in Fisher's (1930) Fundamental Theorem of Natural Selection (FTNS) on "the rate of increase in fitness of any organism at any time" (p. 35) understood as "the rate of increase in the mean value [...] produced by Natural Selection" (p. 42) is that "any increase dp in the proportion of one type of gene at the expense of the other will be accompanied by an increase αdp in the average fitness of the species" (p. 34). This increase is corrected to $2\alpha dp$ for a diploid population in Fisher (1941) whose main objective is to clarify the meaning of the average effect α. The fact that this increase corresponds to Wdt where W is the genetic variance in fitness (called today additive genetic variance) gives the rate of increase in fitness meant in the FTNS.

The average effect of a given gene substitution on any measurement is defined as follows in Fisher (1941, p. 53–54): "It is natural to conceive this as the actual increase in the total of the measurements of a population, when without change in the environment, or in the mating system, the gene substitution is experimentally brought about, as it might be by mutation. [...] To find the effect of the gene substitution in a constant environment, any effect of the gene substitution itself on the environment must be discounted. In other words, the effect that is wanted is only that due to the change in the frequencies of the different possible genotypes, not including any change in the average measurement of a given genotype, which the change in gene frequency may in fact bring about."

In a context of two alleles at one locus, the above verbal definition is followed by a mathematical definition (p. 54) : "The direct mathematical measure of the average effect of a proposed gene substitution is the partial regression, in the population as actually constituted, of the genotype measurement on the numbers 0, 1 or 2 of the allelomorphic genes in each genotype."

Therefore, the calculation of the average effect of substituting G for g, represented by α, is made by assigning genetic values $\mu+\alpha$, μ and $\mu-\alpha$ to the genotypes GG, Gg and gg associated with measurements i, j and k and occurring with frequencies P, $2Q$ and R, respectively, such that the sum of squares

$$P(i-(\mu+\alpha))^2+2Q(j-\mu)^2+R(k-(\mu-\alpha))^2 \tag{9}$$

is minimum. The solution is found to be

$$\alpha=\frac{P(Q+R)}{P(Q+R)+R(P+Q)}(i-j)+\frac{R(P+Q)}{P(Q+R)+R(P+Q)}(j-k). \tag{10}$$

This means that α is the average change in the genotypic measurement for one substitution of G for g with this substitution occurring on genotype Gg or gg with probability

$$u=\frac{P(Q+R)}{P(Q+R)+R(P+Q)} \tag{11}$$

or

$$1-u=\frac{R(P+Q)}{P(Q+R)+R(P+Q)}, \tag{12}$$

respectively. Moreover, it is noted that the change in the total measurement of the population

$$idP+2jdQ+kdR=(i-j)dP-(j-k)dR \tag{13}$$

is ascribable to a change dp in the frequency of G and given by $2\alpha dp$, if the change in genotype frequencies is given by

$$dP=2udp=\frac{2P(Q+R)}{P(Q+R)+R(P+Q)}\,dp, \tag{14}$$

$$dR=-2(1-u)dp=-\frac{2R(P+Q)}{P(Q+R)+R(P+Q)}\,dp \tag{15}$$

and

$$2dQ=-dP-dR=2(1-2u)dp=\frac{2Q(R-P)}{P(Q+R)+R(P+Q)}\,dp. \tag{16}$$

These expressions are the formulas given in Fisher (1941, p. 55) times a missing common factor 2 in order to have $dP+dQ=dp$. Furthermore, they entail the relationship

$$\frac{dP}{P}+\frac{dR}{R}=\frac{2dQ}{Q}, \tag{17}$$

which means that "Q^2/PR maintains a constant ratio λ" as stated in Fisher (1941, p. 56), since then

$$d\left(\frac{Q^2}{PR}\right) = \frac{2PRQdQ - Q^2(PdR + RdP)}{P^2R^2} = 0. \tag{18}$$

More importantly, this condition is used to define a constant genic environment, and it is implied for a change in genotype frequencies to be ascribable only to change in gene ratio, which is equivalent to change in gene frequencies in the case of two alleles.

In Fisher (1941, p. 57), it is recalled that the rate of increase in the average fitness ascribable to natural selection meant in the FTNS corresponds to the increase of average fitness ascribable to a change in gene frequency dp given by $2\alpha dp$, actually ascribable only to a change in gene frequency dp owing to what precedes with fitness as measurement.

Therefore, we are in the following situation. The increase in fitness produced by Natural Selection in Fisher's (1930) FTNS is defined as the increase in the average fitness ascribable only to change in gene ratio but this increase is ascribable only to change in gene ratio because it is given by the proposed increase. Consequently, saying that the variance of the genetic values, called today the additive genetic variance, is the rate of increase ascribable only to change in gene ratio, or produced by Natural Selection in this sense, is an arbitrary statement.

Note that the condition for a constant genic environment given in Eq. (17) ensures that the residual addend vanishes even in the non-additive case, that is,

$$\begin{aligned}(i-(\mu+\alpha))dP + 2(j-\mu)dQ + (k-(\mu-\alpha))dR \\ = (i-j-\alpha)dP - (j-k-\alpha)dR = 0,\end{aligned} \tag{19}$$

not only when $i-j = j-k = \alpha$, but also when $i-j \neq j-k$. Moreover, for the change in genotype frequencies to depend only on the change in gene ratio as in (14)–(16), the genotype frequencies have to be functions only of the gene ratio. Note that the average effect of substituting G for g with the relative changes in genotype frequencies dP/P, dQ/Q and dR/R as measurements for GG, Gg and gg, respectively, is given by

$$\alpha = \frac{1}{P(Q+R) + R(P+Q)}\, d(P+Q). \tag{20}$$

This uses the fact that $P+2Q+R=1$. This fact implies also that the mean measurement is 0 and, therefore, that the genetic values of the measurements can be expressed in the form $2\beta_1$, $\beta_1+\beta_2$ and $2\beta_2$, respectively, with $\beta_1 = (Q+R)\alpha$ and $\beta_2 = -(P+Q)\alpha$. With the fact that $P+Q=p$, this means that the change in genotype frequencies given in (14)–(16) is actually a partial change obtained from a least-squares linear approximation of the relative changes in genotype frequencies as determined by the alleles present in each genotype. Hence, the main property of the average effect of any measurement is actually explained by the average effect of another measurement.

In Fisher (1958), the average effects of substituting given alleles for a random selection of the alleles present at the same locus are considered in a context of multiple alleles, but it is the same apparently circular reasoning and the same assumption on genotype frequencies that lead to the statement of the FTNS.

As for the genetic variance in fitness, which is the largest possible component of variance explainable by gene frequencies and known today as additive genetic variance, it can be said at best that it corresponds to a partial change in mean fitness (Price, 1972; Ewens, 1989; Lessard, 1997; Edwards, 2002; Ewens and Lessard, 2015).

Appendix B. Conditions for a constant genic environment

The conditions for a constant genic environment may not have received sufficient attention in the literature, notable exceptions being Kempthorne (1957), Kimura (1958), Lessard (1997) and Edwards (2002, 2014), and we therefore discuss it here. To do so we consider first the case where the fitness of any individual depends only on the genes at some locus "A". Suppose that only two alleles, A_1 and A_2, are possible at this locus. Let P_{11}, $2P_{12}$ and P_{22} be the respective frequencies of the genotypes A_1A_1, A_1A_2 and A_2A_2 at the time of conception of any parental generation and ΔP_{11}, $2\Delta P_{12}$ and ΔP_{22} be the changes in these genotype frequencies between that time and the time of conception of the corresponding offspring generation. Fisher (1941) in effect considered in some detail the case when the relative changes in genotype frequencies satisfy the linear equation

$$\frac{\Delta P_{11}}{P_{11}} + \frac{\Delta P_{22}}{P_{22}} = \frac{2\Delta P_{12}}{P_{12}}, \tag{21}$$

and referred to this situation as one of "constant genic environment" (see Appendix A for the original continuous-time setting). More generally one can consider the case where k alleles are possible at this locus and the linear equation

$$\frac{\Delta P_{ii}}{P_{ii}} + \frac{\Delta P_{jj}}{P_{jj}} = \frac{2\Delta P_{ij}}{P_{ij}} \tag{22}$$

holds for every (i, j) pair with $i \neq j$. Ewens (2004), following Lessard (1997), showed that when all possible equations of the form (22) hold, the total change in mean fitness from one generation to the next is equal to a partial change in mean fitness in agreement with Fisher's statement of the FTNS. It is therefore relevant to investigate the circumstances under which equations of the form (22) hold, and to assess what implications these equations have if they do hold.

To a first-order approximation, Eq. (22) holds, as was in effect stated by Fisher (1941), if $P_{ij}^2/(P_{ii}P_{jj})$ is unchanging over time for every (i, j) with $i \neq j$, and the most realistic case when this holds is when Hardy-Weinberg frequencies obtain, or in effect when mating is at random. This then recovers the classroom result that when mating is at random and fitnesses depend on the genes at one locus only, mean population fitness increases over time (or remains constant), and that when all genotypic fitnesses are close to each other, the increase in mean population fitness from one generation to the next is close to the parental generation additive genetic variance in fitness. However, the random mating requirement does not seem to correspond to a reasonable concept of constant genic environment. Moreover, random mating does not exactly imply Eq. (22) for a discrete-time model with non-overlapping generations, and does not at all imply the corresponding linear equation

$$\frac{dP_{ii}}{P_{ii}} + \frac{dP_{jj}}{P_{jj}} = \frac{2dP_{ij}}{P_{ij}} \tag{23}$$

for the relative changes in genotype frequencies for $i \neq j$ in a continuous-time model with overlapping generations, since then Hardy-Weinberg proportions are generally approached only in the long run even under neutrality.

A biological interpretation of Eq. (22) or its continuous-time analogue (23) is puzzling. Let us recall that they come from conditions on changes in genotype frequencies for the change in a mean measurement to be given by average effects times changes in gene frequencies (see Appendix A). First, these conditions do not necessarily have a biological interpretation. The statistical notion of average effect can be defined more generally for pairs of any objects and the probability distribution of pairs can change according to rules that have nothing to do with genetics. Second, changes in genotype frequencies do not generally depend only on

changes in gene frequencies as required. This is the case, however, for partial changes obtained by replacing the relative changes in genotype frequencies with their additive genetic values. These changes have been said ascribable only to changes in gene frequencies (Fisher, 1941). But they do not generally take into account all changes in gene frequencies.

In the limit of weak selection with any regular system of mating as partial selfing, for instance, the genotype frequencies at one locus with two possible alleles A_1 and A_2 are given by

$$\begin{aligned} P_{11} &= Fp + (1-F)p^2, \\ P_{22} &= F(1-p) + (1-F)(1-p)^2, \\ 2P_{12} &= (1-F)2p(1-p), \end{aligned} \tag{24}$$

where $p = P_{11} + P_{12}$ is the frequency of A_1 and F is the inbreeding coefficient (with partial selfing, $F = s/(2-s)$ where s is the probability of selfing). Then the total changes in genotype frequencies in continuous time are

$$\begin{aligned} dP_{11} &= (F + 2p(1-F))dp, \\ dP_{22} &= -(F + 2(1-p)(1-F))dp, \\ 2dP_{12} &= 2(1-2p)(1-F)dp, \end{aligned} \tag{25}$$

while the changes said ascribable to a change dp (in a constant genic environment) given in Fisher (1941) (see Appendix A with $P = P_{11}, R = P_{22}$ and $Q = P_{12}$) are given in this case (after simple algebraic manipulations) by

$$\begin{aligned} dP_{11} &= \frac{2(F + p(1-F))}{1+F}\, dp, \\ dP_{22} &= -\frac{2(F + (1-p)(1-F))}{1+F}\, dp, \\ 2dP_{12} &= \frac{2(1-2p)(1-F)}{1+F}\, dp. \end{aligned} \tag{26}$$

These changes are different from the total changes due to a change dp in the frequency of A_1 unless $F = 0$, which means genotypes in Hardy-Weinberg proportions.

Fisher (1930, 1958) actually never considered the whole-genome analysis associated with Eq. (22). It is known (Ewens, 2004, p. 260) that even if these equations hold for all possible pairs of alleles at all possible loci in the genome, it is no longer true in the whole-genome case that the total change in mean fitness from one generation to the next is equal to the partial change in mean fitness unless there is total linkage equilibrium. Actually, the partial change in the whole-genome case with constant fitnesses is equal to the total change if the relative change in the frequency of genotype g, namely $\Delta P_g/P_g$ in discrete time or dP_g/P_g in continuous time, is linear with respect to gene frequencies in g (Lessard, 1997), which extends conditions (22) and (23). It is therefore no longer necessarily true in the whole-genome case, even under random mating, that population mean fitness increases (or remains constant) from one generation to the next as has been known for a long time since an exact analysis in the two-locus case (Moran, 1964), in contrast to the situation for the one-locus case. It is known, however, that if selection is weak in a random mating population undergoing discrete, non-overlapping generations, then the population evolves approximately as if it were in linkage equilibrium and the total change in mean fitness from one generation to the next is approximately given by the additive genetic variance in fitness divided by the mean fitness (Nagylaki, 1993).

References

Crow, J.F., Kimura, M., 1970. An Introduction to Population Genetics Theory. Harper and Row, New York, NY.

Edwards, A.W.F., 1994. The fundamental theorem of natural selection. Biol. Rev. 69, 443–474.

Edwards, A.W.F., 2002. The fundamental theorem of natural selection. Theor. Popul. Biol. 61, 335–337.

Edwards, A.W.F., 2014. R.A. Fisher's gene-centred view of evolution and the fundamental theorem of natural selection. Biol. Rev. 89, 135–147.

Ewens, W.J., 1989. An interpretation and proof of the fundamental theorem of natural selection. Theor. Popul. Biol. 36, 167–190.

Ewens, W.J., 2004. Mathematical Population Genetics I. Theoretical Introduction, Second Edition. Springer-Verlag, New York, NY.

Ewens, W.J., Lessard, S., 2015. On the interpretation and relevance of the fundamental theorem of natural selection. Theor. Popul. Biol. 104, 59–67.

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

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 Edition. Dover, New York, NY.

Frank, S., Slatkin, M., 1992. Fisher's fundamental theorem of natural selection. Trends Ecol. Evol. 7, 92–95.

Grafen, A., 2015a. Biological fitness and the fundamental theorem of natural selection. Am. Naturalist 186, 1–14.

Grafen, A., 2015b. Biological fitness and the Price equation in class-structured populations. J. Theor. Biol. 373, 62–72.

Grafen, A., 2018. The left hand side of the fundamental theorem of natural selection. J. Theor. Biol. 456, 175–189.

Hamilton, W.D., 1964. The genetical evolution of social behaviour I. J. Theor. Biol. 7, 1–16.

Kempthorne, O., 1957. An Introduction to Genetic Statistics. Wiley, New York, NY.

Kimura, M., 1958. On the change of population fitness by natural selection. Heredity 12, 145–167.

Lee, J.J., Chow, C.C., 2013. The causal meaning of Fisher's average effect. Genet. Res. 95, 89–109.

Lessard, S., 1997. Fisher's fundamental theorem of natural selection revisited. Theor. Popul. Biol. 52, 119–136.

Lessard, S., Soares, C., 2016. Definitions of fitness in age-structured populations: comparison in the haploid case. J. Theor. Biol. 391, 65–73.

Lessard, S., Soares, C.D., 2018. Frequency-dependent growth in class-structured populations: continuous dynamics in the limit of weak selection. J. Math. Biol. 77, 229–259.

Moran, P.A.P., 1964. On the nonexistence of adaptive topographies. Ann. Hum. Genet. 27, 383–393.

Nagylaki, T., 1976. The evolution of one- and two-locus systems. Genetics 83, 583–600.

Nagylaki, T., 1993. The evolution of multilocus-locus systems under weak selection. Genetics 134, 627–647.

Okasha, S., 2018. Agents and Goals in Evolution. Oxford University Press, Oxford, UK.

Price, G.R., 1972. Fisher's 'fundamental theorem' made clear. Ann. Hum. Genet. 36, 129–140.

Walsh, B., Lynch, M., 2018. Evolution and Selection of Quantitative Traits. Oxford University Press, Oxford, UK.