Brit. J. Phil. Sci. **67** (2016), 247–269

Causal Foundations of Evolutionary Genetics

Jun Otsuka

ABSTRACT

The causal nature of evolution is one of the central topics in the philosophy of biology. The issue concerns whether equations used in evolutionary genetics point to some causal processes or purely phenomenological patterns. To address this question the present article builds well-defined causal models that underlie standard equations in evolutionary genetics. These models are based on minimal and biologically plausible hypotheses about selection and reproduction, and generate statistics to predict evolutionary changes. The causal reconstruction of the evolutionary principles shows adaptive evolution as a genuine causal process, where fitness and selection are both causes of evolution.

1 Introduction

The causal nature of evolution is one of the central topics in the philosophy of biology. Is evolution a causal process? Are selection and fitness causes of

doi:10.1093/bjps/axu039

Advance Access published on December 4, 2014

population change? Recent discussions in the literature have given conflicting answers to these questions. The causal scepticism is motivated by the fact that most, if not all, principles of evolutionary theory—such as the Price equation or Fisher's fundamental theorem of natural selection—are expressed by purely statistical terms such as variances or covariances. This does not preclude, however, the possibility that such statistics are products of certain causal structures. The past few decades have seen the development of a mathematical framework for studying causal relations and the probability distributions generated by them (Pearl [2000]; Spirtes *et al*. [2000]). Using this technique, the present article shows how the Price equation—one of the 'fundamental' principles of evolutionary theory—is generated from causal relationships representing selection and reproduction. The derivation provides causal foundations for the standard equations used in evolutionary genetics and establishes adaptive evolution as a *bona fide* causal process.

The structure of the article is as follows: After a brief description of the problem in Section 2, basic notions of causal models are introduced in Section 3. Using this machinery, Section 4 investigates causal models underlying the Price equation, the breeder's equation in quantitative genetics, and the one-locus population genetics system. The resulting models are causal in the sense that they can be used to predict the consequence of an ideal intervention, and evolutionary to the extent that they describe or predict changes in population frequencies induced by selection. The explicit definition of causal structures brings several philosophical upshots (Section 5). The first corollary is that selection must be understood as a causal process (a trait affecting fitness), rather than just an outcome (statistical dependence between the trait and fitness). Second, the causal models give clear-cut answers to the entangled questions of whether fitness and/or selection cause population change. Applying the formal intervention calculus (Pearl [2000]; Spirtes *et al*. [2000]) to the causal models obtained in Section 4, I will show there are some interventions on selection and fitness that affect evolutionary outcomes. This result gives an unequivocal proof that fitness and selection are both genuine causes of evolution.

One disclaimer before proceeding: this article exclusively focuses on selection, leaving aside other factors of evolutionary changes such as mutation, migration, or drift. Exclusion of drift means all distributions should be taken as population distributions in infinite populations. The causal basis for drift may be treated on another occasion.

2 The Philosophical Puzzle

Modern mathematical theories of evolution study changes in populations by state transition functions (Lewontin [1974]; Lloyd [1988]). Such functions

describe a temporal change in certain features of a population based on its current state, thereby allowing a prediction of its evolutionary trajectory. Taking the simplest example, in the two allelic system with no dominance, where the fitnesses of genotype AA, Aa, and aa are, respectively, $1 + s$, $1 + s/2$, and 1, the change in the population frequency, p, of alleles A between two consecutive generations is given by:[1]

$$\Delta p = \frac{sp(1-p)}{2(sp+1)}. \tag{1}$$

Alternatively, we may be interested in the evolutionary change of a phenotype rather than of a gene. In quantitative genetics, the between-generation change in the phenotypic mean $\overline{Z}$ is given by the breeder's equation:

$$\Delta\overline{Z} = h^2 S, \tag{2}$$

where the selection differential, S, is measures the shift in the phenotypic mean of the parental generation after selection (but before reproduction) and the heritability, h^2, denotes the efficiency of reproduction—that is, how much of the change induced by selection is passed onto the next generation. Note that both Equation (1) and Equation (2) describe the dynamics of population features (a change in genetic frequencies or the phenotypic mean) based on certain characteristics of the current population. The importance of such state transition functions cannot be overstated, for it is these mathematical analyses that integrated Mendelian genetics and Darwin's theory of natural selection, and formed the core part of evolutionary theory after the Modern Synthesis.

On the other hand, however, the algebraic treatments obscure the question of evolutionary causes. Although successful state transition functions may give correct or at least acceptable predictions of a future population under certain conditions, this does not automatically mean they represent causal processes. For, obviously, transition functions may describe non-causal as well as causal patterns. One could write down sufficiently predictive models for planetary motion without knowing the Newtonian mechanics (for example, Kepler's laws). Such models are purely phenomenological, rather than causal. Thus philosophers have long been concerned with whether the evolutionary equations mentioned above point to any causal process, and if so, how.

Elliott Sober ([1984]), for example, argues that the causal contents of an evolutionary equation are furnished by the 'source laws', which estimate or measure parameters and variables in the equation by empirical means such as functional analysis. Since such estimates should reflect causal facts about organisms, Sober claims the whole of evolutionary theory is causal and

[1] Throughout this article I assume generations to be discrete and non-overlapping.

empirical, even if its core equations—what he calls the 'consequence laws'—may be purely mathematical.

This view has been vigorously challenged recently by a group of philosophers called statisticalists (Matthen and Ariew [2002], [2005], [2009]; Walsh *et al.* [2002]; Walsh [2007], [2010]; Matthen [2010]), who argue that the quantities appearing in evolutionary equations, especially fitness, cannot be estimated by Sober's source laws or any other causal analysis of a similar kind, but only by census. Decoupling evolutionary equations from underlying mechanisms, the statisticalists insist that modern genetics gives a purely phenomenological description of the statistical trends of a population, or in their words, 'explains the changes in the statistical structure of a population by appeal to statistical phenomena' (Walsh *et al.* [2002], p. 471).

Parallel to this issue—whether evolutionary theory describes a causal process or not—is the question of whether its key concepts, most notably fitness and selection, identify causes of evolutionary change. Millstein ([2006]), for example, argues that selection is a population-level cause of evolution, while Matthen and Ariew ([2009]) and Lewens ([2010]) deny any causal power to selection. Walsh ([2007], [2010]) claims that fitness is causally inert since it fails to satisfy certain criteria of causality, while this argument was criticized by Otsuka *et al.* ([2011]). Sober ([2013]), finally, argues that fitness itself does not cause population change, but its variance does. Although the reasonings behind these claims vary, they all base at least some portion of their argument on the manipulationist notion of causation (Woodward [2003]). That is, both parties seem to agree that fitness and selection are causal if and only if manipulating them affects evolutionary response.

But how do we know the consequence of such manipulations? To examine this, most (but not all) of these authors resort to conceptual analysis: what really are fitness and selection? What do they stand for? With a certain interpretation of these concepts, they go on to argue that the supposed manipulation should (or should not) affect evolution, and thus that the concepts must (or must not) be causal.

This, to say the least, is a very peculiar move. In the manipulationist framework, the outcome of a possible intervention is not determined by the meaning of variables, but their relationships. Woodward's example (Woodward [2003], p. 197) makes it clear: It is known that the period, T, of a simple pendulum is related to its length, l, by

$$T = 2\pi\sqrt{l/g}, \tag{3}$$

where g is the acceleration due to gravity. It seems natural to read this equation causally, to the effect that the right-hand side (the length and gravity)

determines or 'causes' the left hand side (the period), until we find that Equation (3) is mathematically equivalent to the following:

$$l = \frac{T^2 g}{4\pi^2}. \tag{4}$$

Now it is obviously absurd to claim, based on this new equation, that the period causes the length of the pendulum. What determines the (il)legitimacy of the causal reading of each equation? Surely not the meanings of the variables, for they stay the same in both equations.

The moral of this simple example is that the conceptual analysis is utterly irrelevant to the investigation of the causal nature of some concept under the manipulationist framework. Sober ([2013]) resorts to the breeder's equation (Equation (2)) to make his case that the fitness variation, measured by the term S on the right-hand side, affects the response to selection in the left-hand side. But this begs the question. How do we know the breeder's equation correctly captures the causal flow? Why isn't it like Equation (4), rather than Equation (3)?[2] We never know, until the causal relationships among the variables are explicitly specified beforehand. Such relations are usually given by a causal model, which also determines a set of equations that allow for causal reading (Pearl [2000]; Spirtes *et al.* [2000]). It is the causal model given by the Newtonian mechanics that authorizes the causal reading of Equation (3), but not of Equation (4). In the same way, if we want to know the effect of intervening on some variable in an evolutionary formula, we need the causal model underlying that equation. Hence the second contention—whether fitness or selection causes evolution—hinges on the first: are there causal models underlying evolutionary transition functions?

The answer is yes. This article describes causal models that (1) include the relevant variables such as genetic, phenotypic, and environmental factors, (2) generate the statistics necessary to describe and predict evolutionary trajectories, and (3) can be used to predict the consequence of a possible intervention on an evolving population. The derived models will reveal the causal foundations underlying the evolutionary transition functions as described above (Equations (1) and (2)), and help us to determine whether fitness and selection can be properly regarded as causes of evolutionary change.

[2] In the breeder's equation, the evolutionary response cannot precede fitness variance, and hence cannot be its cause. But they may be effects of a common cause, or, as statisticalists may argue, the relation may be 'purely statistical'. In fact, we will later see that, *pace* Sober, a manipulation of the fitness variance does not affect the expected evolutionary response predicted by the breeder's equation.

3 Causal Models

A causal model employs a graphical structure to represent causal relationships among variables (Pearl [2000]; Spirtes *et al*. [2000]). A causal graph $\mathcal{G} = (\mathbf{V}, \mathbf{E})$ is a pair comprising a set of variables $\mathbf{V}$ (or nodes) and a set of edges $\mathbf{E} \subseteq \mathbf{V} \times \mathbf{V}$. An edge $(X, Y) \in \mathbf{E}$, or more graphically $X \to Y$, represents a direct causal relation from X to Y, where X is called a parent of Y and Y a child of X. A path between X and Y is any chain of edges between X and Y, where a path can follow arrows in either the direction of the arrow or the reverse direction. If every arrow in a path between X and Y is pointing towards Y, it is called a directed path from X to Y, and then X is a cause of Y and Y is an effect of X. A bidirected edge $X \leftrightarrow Y$ represents unmodelled association between X and Y—that is, the association not accounted for by any causal path in the graph. In this article, such edges are allowed only between those variables having no causes/parents (called exogenous).

It is assumed that the value of each variable $V_j \in \mathbf{V}$ is determined by its direct causes or parents $\mathbf{PA}(V_j)$ such that

$$V_j = f_j(\mathbf{PA}(V_j)). \tag{5}$$

This is called the structural equation for V_j. When the relationship is linear, as assumed throughout this article, Equation (5) can be expressed as

$$V_j = \sum_{V_i \in \mathbf{PA}(V_j)} \beta_{ji} V_i, \tag{6}$$

with a set of linear coefficients $\boldsymbol{\beta}$ (also called path coefficients). Hence, in a linear causal model each directed edge in the graph is associated with one linear coefficient.

A causal graph $\mathcal{G}$ over $\mathbf{V}$, a set of corresponding structural equations $\mathbf{F}$, and a probability distribution P over the exogenous variables in $\mathbf{V}$ uniquely determine the joint distribution over $\mathbf{V}$. The induced distribution satisfies useful properties such as the Markov Condition (Pearl [1988]). Another feature of our interest is the trek rule (Wright [1921]). A trek between variables X and Y is a path between them that does not contain a collider where two arrows on the path collide at one variable (for example, $\to V \leftarrow$). A trek is equivalent to a pair of directed paths that share the same source or whose separate sources are connected by a bidirected edge.[3] Thus in Figure 1, $X_1 \to X_3 \to X_5$, $X_3 \leftarrow X_2 \to X_4$, and $X_5 \leftarrow X_3 \leftarrow X_1 \leftrightarrow X_2 \to X_4 \to X_5$ are examples of treks, whereas $X_3 \to X_5 \leftarrow X_4$ is not. For each trek, we can calculate the trek coefficient by multiplying the (co)variance of its source(s) and all the linear coefficients on the edges constituting the trek. The trek rule states that

[3] Note that one of the pair may be empty. Thus one directed path from X to Y counts as a trek between them.

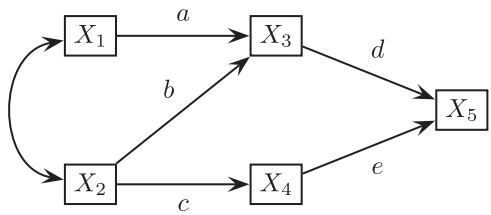

Figure 1. An example of a causal graph with path coefficients.

the covariance of two variables equals the sum of trek coefficients over all the treks connecting them. That is, if **T** is the set of all the treks between X and Y, and β_{ti} is the linear coefficient of the ith edge on $t \in \mathbf{T}$,

$$\mathrm{Cov}(X, Y) = \sum_{t \in \mathbf{T}} \sigma_t \prod_{i \in t} \beta_{ti}, \tag{7}$$

where σ_t is the (co)variance of the source(s) of trek t. To take some examples from Figure 1, $\mathrm{Cov}(X_1, X_4) = \mathrm{Cov}(X_1, X_2)c$, $\mathrm{Cov}(X_3, X_4) = \mathrm{Cov}(X_1, X_2)ac + \mathrm{Var}(X_2)bc$, and $\mathrm{Cov}(X_3, X_5) = \mathrm{Var}(X_3)d + \mathrm{Cov}(X_1, X_2)ace + \mathrm{Var}(X_2)bce$.

Causal models thus give formal tools to study the relationships between a causal structure and the probability distribution generated by it. The next section makes use of these tools to explore causal bases of evolutionary equations.

4 Causal Foundations of Evolutionary Genetics

In evolutionary genetics, it is well known that a change in moments (for example, mean) of a population from one generation to the next is completely described by the Price equation (Robertson [1966]; Price [1970]). Let Z be the trait of interest, W be the Darwinian fitness as defined by the number of offspring, and Z' be the average phenotype of offspring of each individual. Thus if George, who reproduces asexually, has four children each having the phenotypic value of 1, 1, 1, and 2, then $w_{\mathrm{George}} = 4$ and $z'_{\mathrm{George}} = (1 + 1 + 1 + 2)/4 = 1.25$.[4] The Price equation gives the difference, $\Delta\overline{Z}$, of mean phenotypic values between the parental generation and the offspring generation by

$$\Delta\overline{Z} = \frac{1}{\overline{W}}\mathrm{Cov}(W, Z') + \overline{Z}' - \overline{Z}, \tag{8}$$

where the upper bars denote the means. The first term of the equation is the covariance of the fitness and the averaged offspring phenotype, and thus reflects both selection and reproduction. The second and third terms, in contrast, compare the phenotype of parents and the averaged phenotypic value of their offspring, regardless of the fitness of the parents. A difference in these

[4] Throughout this article, random variables are denoted by uppercase letters while their values are denoted by lowercase letters. Boldface is used for sets, vectors, and matrices.

terms, therefore, implies a transmission bias.[5] In this article I will assume transmission bias to be absent, in which cases evolutionary dynamics is described just by the first term.

Before moving on, let us emphasize that the variables used in the Price equation, including fitness W, are all properties of an individual (or of a pair of individuals for diploid organisms, as we will see later). Alternatively the concept of fitness is sometimes used to refer to a property of types, for example, phenotype, genotype, haplotype, or an allele. Such type-level fitnesses are called marginal fitness and defined by the conditional distribution $P(W|T = t)$ or its mean for a given type, t. But what we denote by 'fitness' in this article is primarily a property of an individual.[6] The Price equation thus gives population change, $\Delta\overline{Z}$, as a statistical function of these individual variables.

A remarkable feature of the Price equation is that it is a mathematical theorem and thus holds true of any evolving population. This has motivated the view that the core evolutionary principles are *a priori* truths (for example, Sober [1993], p. 72) and at the same time generated the philosophical puzzle as to how such non-empirical theorems can represent causal processes in the real world. Indeed, Price's theorem does not tell us how the variables in the equation affect each other or what will happen if one of them is altered by some external means—or in other words, it does not explain why evolution takes place in that way. As we saw in Section 2, answering these questions requires a suitable causal model beyond a mere mathematical equation.

The goal of this section is to find such causal foundations for evolutionary change represented by the Price equation. Our strategy is to build causal models (that is, specify causal graphs and structural equations) representing evolutionary processes and then show that such models indeed generate the Price covariance, $\mathrm{Cov}(W, Z')$. This will give us evolutionary state transition functions that have a definite causal basis and describe evolutionary changes in terms of causal parameters. I will show this for phenotypic evolution first, and then consider the population genetics model.

4.1 Univariate quantitative genetics model

4.1.1 The causal graph

Evolution by natural selection consists of two parts: selection and reproduction. Let us take reproduction first. Reproduction is a process that links the

[5] This 'transmission bias', however, may include selection at lower levels (such as genic selection for 'selfish genes') and effects of non-genetic inheritance (such as maternal effects).

[6] Some statisticalists (for example, Pigliucci and Kaplan [2006]) seem to interpret fitness to be a population level feature—that is, as a random variable or the expectation thereof defined over a set of populations—but no such use of the concept is warranted by the evolutionary literature. See (De Jong [1994]) for a discussion of various concepts of fitness.

phenotype of parents with the phenotype of offspring through genes or epigenetic materials. Thus a causal model for reproduction must specify how a phenotype is formed out of these factors and how they are transmitted from parent to offspring. Obviously there are many possible reproductive structures, but here we confine ourselves to a very simple case of purely Mendelian inheritance.

Suppose there are n different types of alleles segregating in a population.[7] Then the genotype of an organism is characterized by a set (vector) of n variables $\mathbf{X} := (X_1, X_2, \ldots, X_n)$, where $X_i \in \mathbf{X}$ is the gene content, that is, the count of copies of the ith allele type in an individual (Lynch and Walsh [1998], p. 65). For a haploid organism, the value x_i of X_i for any i can be either 0 or 1, while for diploids $x_i \in \{0, 1, 2\}$.

Parental phenotype Z is made out of these genes as well as of an environmental factor denoted by E_Z. We thus have edges drawn from E_Z and each of $\mathbf{X}$ to Z. These genes are then transmitted to offspring, following the causal edge from parental to offspring gene contents, $X_i \rightarrow X'_i$, for each i. Finally, we assume the same developmental process for offspring phenotype, Z' being caused by $\mathbf{X}'$ and E'_Z.

The above construction gives the causal graph for reproduction as shown in Figure 2 (the path coefficients in the graph will be explained shortly). The graph, however, makes further assumptions not mentioned above. First, bidirected edges between parental genes represent genetic correlations already present in the population. Such correlations can arise in two ways: gene counts of the same locus are necessarily correlated for they must sum up to the ploidy of the organism, while inter-locus correlations, often called 'linkage disequilibrium' or 'gametic phase disequilibrium' arise due to various factors including previous selection, drift, or non-random mating.[8] In contrast, it is assumed that environment E_Z is not correlated with any genes, as implied by the absence of bidirected edges between $X \in \mathbf{X}$ and E_Z. The graph also presupposes that parental environment E_Z has no causal influence on, or correlation with, offspring environment E'_Z. Finally, transmission is strictly Mendelian in the sense that each gene is inherited independently, without affecting the transmission process of other genes—this excludes segregation distortion.

We now move to the second part, selection. Selection is the process in which parental phenotypes lead to differential reproductive success. We say that trait Z is selected if and only if it, along with an environmental factor denoted by

[7] In standard treatments these n allele types are partitioned into a set of loci. The partition becomes necessary in order to distinguish two types of genetic interactions: dominance and epistasis. But here we ignore this because the breeder's equation does not consider any non-linear genetic interactions.

[8] These empirical covariances can be seen as a dependency due to a selection bias.

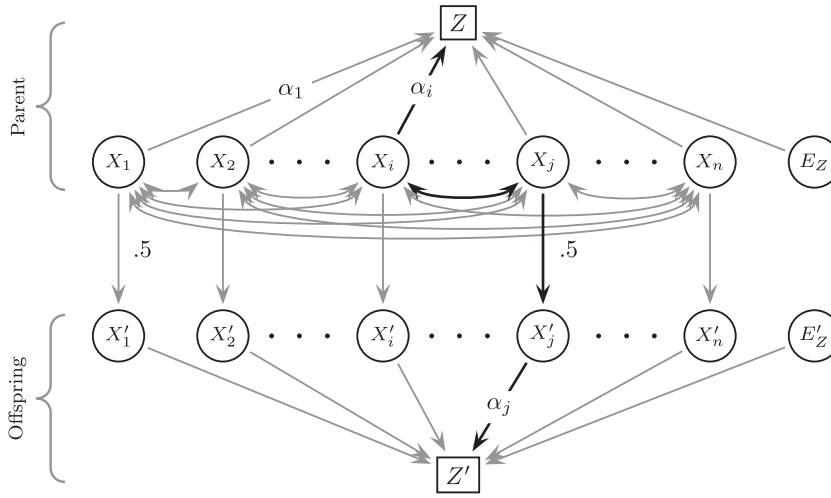

Figure 2. Linear (additive) decomposition of the covariance between parental and offspring traits. Bold arrows illustrate an example of a trek connecting Z and Z', whose contribution to the covariance is $\alpha_i \mathrm{Cov}(X_i, X_j) \cdot \frac{1}{2} \cdot \alpha_j$. See the main text for the explanation of the variables.

E_W, causally affects fitness W (for example, Glymour [2011]).[9] This means that in order for Z to be selected there must be some intervention on Z, at least as a possibility, that changes fitness W. Selection can thus be represented in Figure 2 by adding edges $Z \rightarrow W$ and $E_W \rightarrow W$.

A slight complication arises, however, for diploid organisms that do not produce offspring by themselves but only by a pair. It follows that the proper unit for analysing diploid evolution is a male and female pair. For a given pair, let us denote the phenotypes of the female and the male by Z_F and Z_M, and their gene contents by $\mathbf{X}_F$ and $\mathbf{X}_M$, respectively. Fitness, W, of the pair is the number of offspring produced by that pair, and has Z_F and Z_M (and E_W) as its direct causes. Likewise, Z' and $\mathbf{X}'$ are redefined to be the average phenotypic value and gene contents of the offspring of that pair, respectively.

With these modifications, the overall causal graph that incorporates selection and reproduction of diploid organisms should look like Figure 3, where each branch in the middle (the mother and the father) is an abbreviated representation of the reproductive causal model represented in Figure 2. As before, this graph introduces additional assumptions. First, the environment factors affecting fitness (E_{W_M} and E_{W_F}) must be uncorrelated with phenotypes or genotypes, as implied by the absence of edges between them. Another

[9] I will discuss in Section 5.1 why selection must be defined as a causal process, not just a statistical dependence.

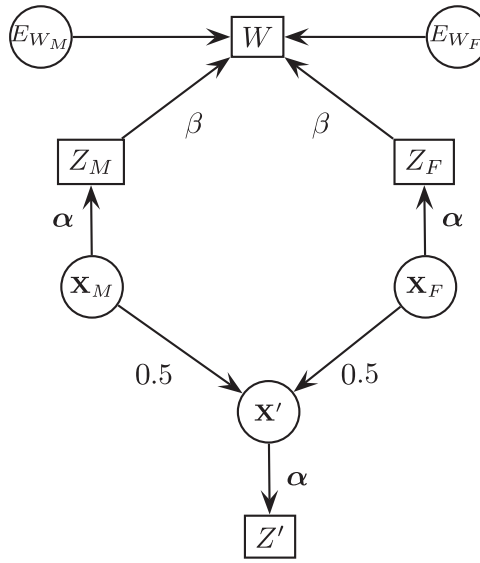

Figure 3. The causal graph showing the connections between parental fitness, W, and offspring trait, Z'. Boldface letters denote multiple nodes or coefficients, for example, $\mathbf{X}_M := (X_{M_1}, X_{M_2}, \ldots, X_{M_n})$, and each side abbreviates the structure shown in Figure 2. Environmental factors for phenotypes E_{Z_M}, E_{Z_F}, and E'_Z are omitted from the graph.

assumption is random mating; non-random mating would introduce bidirected edges between corresponding elements in $\mathbf{X}_M$ and $\mathbf{X}_F$.

Although minimal and even simplistic, Figures 2 and 3 submit a biologically reasonable hypothesis of the causal structure underlying selection and reproduction. It specifies causal links among relevant variables in such a way that we can identify which part of the system would be affected if some of them were manipulated by an external means. It is not yet clear, however, how this causal structure over individual organisms relates to the population changes as described by evolutionary transition functions. To see this relation, we need to quantify each causal relationship appearing in the graph, the task to which we now turn.

4.1.2 Structural equations

Compared to the causal graph, there is much less, if any, *a priori* reason for determining the functional form for a given causal relationship. How a cause contributes to its effects is largely an empirical matter that depends on their nature and circumstance. As a first approximation, however, I assume in this article that every cause acts additively. This means that selection is purely

directional and there is no dominance or epistasis. Non-linear structural equations are possible in theory but complicate the mathematical derivation; more important for our purpose, they are outside the scope of the standard equations of evolutionary genetics mentioned above and this article, since providing the causal structures for these equations is the primary goal of this article.

In linear/directional selection, a unit change in the phenotype affects the fitness by the amount specified by the path coefficient β, so that

$$W = \beta Z + E_W. \tag{9}$$

We further assume the selection pressures to be the same for male and—that is, Z_F and Z_M have the same path coefficient with respect to W.

For the genotype–phenotype mapping, we assume each allele, X_i linearly affects the phenotype by coefficient α_i:

$$\begin{aligned} Z &= \sum_{X_i \in \mathbf{X}} \alpha_i X_i + E_Z \\ &= \boldsymbol{\alpha}\mathbf{X}^{\mathrm{T}} + E_{\mathrm{Z}}, \end{aligned} \tag{10}$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_{\mathrm{n}})$ and T denotes matrix transpose. α_i is called the 'additive effect' and measures the change in the phenotype induced by adding one copy of the ith allele, say from $X_i = 0$ to $X_i = 1$ (Fisher [1930/2006], p. 31). It is assumed that additive effects are the same for all individuals in the population. Hence Equation (10) characterizes the genotype–phenotype mapping of females, males, and offspring.

Under diploid Mendelian inheritance (e.g. no segregation distortion), every gene in a parent has a half-chance to get inherited. Hence, the structural equation representing the genetic transmission is simply

$$\mathbf{X}' = \frac{1}{2}\mathbf{X}. \tag{11}$$

Equations (9) to (11) constitute the structural equations corresponding to the causal graph Figures 2 and 3, as indicated by the path coefficients on edges. Together with the graphs, they tell us how a unit alteration in any variable in the model brings about changes in other parts—that is, they give predictions of effects resulting from a possible intervention. This completes the description of a causal model for a single quantitative trait.

4.1.3 Deriving evolutionary transition functions

The final step employs the trek rule to obtain evolutionary transition functions based on the causal model as defined above. Recall that, according to the trek rule the Price covariance $\mathrm{Cov}(W, Z')$ is given by the sum of trek coefficients between W and Z'. From Figures 2 and 3, each trek connecting W and Z' has

the form either of $W \leftarrow Z \leftarrow X_i \rightarrow X_i' \rightarrow Z'$ or of $W \leftarrow Z \leftarrow X_i \leftrightarrow X_j \rightarrow X_j' \rightarrow Z'$, the trek coefficient of each being $\beta\alpha_i \mathrm{Cov}(X_i, X_j) \cdot \frac{1}{2} \cdot \alpha_j$ for $X_i, X_j \in \mathbf{X}$. Summing all these treks for each side of the two parents yields

$$\begin{aligned} \Delta\overline{Z} &= \frac{1}{\overline{W}} \mathrm{Cov}(W, Z') \\ &= \frac{2}{\overline{W}} \beta \sum_{X_i \in \mathbf{X}} \sum_{X_j \in \mathbf{X}} \alpha_i \mathrm{Cov}(X_i, X_j) \cdot \frac{1}{2} \cdot \alpha_j \\ &= \frac{1}{\overline{W}} \beta \boldsymbol{\alpha} \mathrm{Var}(\mathbf{X}) \boldsymbol{\alpha}^{\mathrm{T}}, \end{aligned} \tag{12}$$

where $\mathrm{Var}(\mathbf{X})$ is the covariances of gene contents and is a function of population genetic frequencies. Hence Equation (12) relates phenotypic change to causal parameters (β and $\boldsymbol{\alpha}$), as well as a distributional feature of the exogenous variables ($\mathrm{Var}(\mathbf{X})$), giving a causal underpinning of evolutionary change.

The same model also reveals the causal basis of the standard formula of quantitative genetics, the breeder's equation (Equation (2)). To see this, let us first derive the additive genetic variance, σ_A^2, which is defined as the part of the phenotypic variance due to the additive effects of gene contents $\mathbf{X}$. Since the variance is nothing but the covariance of a variable with itself, we can apply the trek rule to calculate this value. Noting in Figure 2 that all treks connecting Z to itself have the form $Z \leftarrow X_i \leftrightarrow X_j \rightarrow Z$ with the trek coefficient $\alpha_i \mathrm{Cov}(X_i, X_j) \alpha_j$, the additive genetic variance for Z is

$$\begin{aligned} \sigma_A^2 &= \sum_{X_i \in \mathbf{X}} \sum_{X_j \in \mathbf{X}} \alpha_i \mathrm{Cov}(X_i, X_j) \alpha_j \\ &= \boldsymbol{\alpha} \mathrm{Var}(\mathbf{X}) \boldsymbol{\alpha}^{\mathrm{T}}. \end{aligned} \tag{13}$$

Plugging this into Equation (12) yields

$$\Delta\overline{Z} = \frac{1}{\overline{W}} \beta \sigma_A^2. \tag{14}$$

From standard regression theory, the least squares estimate of linear coefficient β is $\mathrm{Cov}(W, Z)/\mathrm{Var}(Z)$. Letting $\widetilde{W} := W/\overline{W}$ denote the relative fitness, we get

$$\begin{aligned} \Delta\overline{Z} &= \frac{1}{\overline{W}} \frac{\mathrm{Cov}(W, Z)}{\mathrm{Var}(Z)} \sigma_A^2 \\ &= \mathrm{Cov}(\widetilde{W}, Z) \frac{\sigma_A^2}{\mathrm{Var}(Z)} \\ &= S h^2, \end{aligned} \tag{15}$$

where $S := \mathrm{Cov}(\widetilde{W}, Z)$ is the selection differential and $h^2 := \sigma_A^2/\mathrm{Var}(Z)$ is the (narrow-sense) 'heritability'. The breeder's Equation (2) (and Equation (15)),

therefore, is an estimate of the linear evolutionary response generated by the causal structure in Figures 2 and 3. We can thus conclude that the graph and model specified above represent the causal foundation of the standard evolutionary formula in quantitative genetics.[10]

4.2 One-locus population genetics model

The same method can be used to build the causal model for the simple population genetics model as in Equation (1), if one thinks of a gene as a kind of phenotype. Let A and a be two alleles segregating at one locus with the allelic frequencies p and $1-p$, respectively. Gene contents X_1 and X_2 then are counts of allele(s) of A and a in an organism. Let us define our 'phenotype' Z to be the frequency of allele A in one organism. Hence, for diploid organisms $Z = X_1/2$ and its value can be either 0, 0.5, or 1. Noting that the population frequency of allele A equals $\overline{Z}$, its change is given by the Price equation:

$$\Delta p = \Delta\overline{Z} = \frac{1}{\overline{W}}\mathrm{Cov}(W, Z'). \tag{16}$$

Here again we ignore the transmission bias and assume that genes are passed to offspring more or less directly.

The causal graph connecting the relevant variables for a pair of organisms is shown in Figure 4. The non-directed edges in the graph represent the unit conversion between the gene count (X) and the gene frequency (Z) in an individual. Since these two variables point to the same thing, the causal flows remain undisrupted and the trek rule is still applicable. In the graph there are only two treks connecting W and Z', that is, $W \leftarrow Z_M - X_{M1} \rightarrow X'_1 \rightarrow Z'$ and $W \leftarrow Z_F - X_{F1} \rightarrow X'_1 \rightarrow Z'$. Assuming selection acts on each sex equally, that is, $s_F = s_M = s$, the trek sum is

$$\begin{aligned}\frac{1}{\overline{W}}\mathrm{Cov}(W, Z') &= \frac{1}{4\overline{W}}s\mathrm{Var}(X_1)\\ &= \frac{sp(1-p)}{2\overline{W}},\end{aligned} \tag{17}$$

where the second line follows from the fact that the variance of the multinomial random variable X_1 is $2p(1-p)$. Under no dominance, the mean fitness $\overline{W}$ is $p^2(1+s) + 2p(1-p)(1+s/2) + (1-p)^2 = sp + 1$, giving

$$\Delta p = \frac{1}{\overline{W}}\mathrm{Cov}(W, Z') = \frac{sp(1-p)}{2(sp+1)}, \tag{18}$$

[10] In the same fashion one can derive the multivariate version of the breeder's equation—the 'Lande equation' (Lande [1979])—which plays the central role in today's quantitative genetics (Otsuka [2014]).

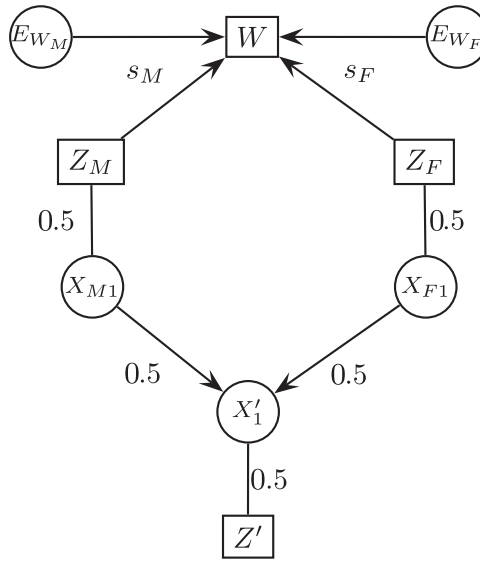

Figure 4. The causal graph for the one-locus population genetics system. Non-directed edges represent mathematical relations (change of units). Variable X_2 is omitted since the gene content of allele a does not affect Zs.

which accords with Equation (1). In general, plugging regression estimate $\hat{s} = \mathrm{Cov}(W, Z)/\mathrm{Var}(Z)$ into Equation (17) yields the standard one-locus population genetics model (Gillespie [2004], p. 62):

$$\Delta p = \frac{p(1-p)[p(w_{AA} - w_{Aa}) + (1-p)(w_{Aa} - w_{aa})]}{\overline{W}}, \tag{19}$$

where w_{AA}, w_{Aa}, and w_{aa} are the marginal fitnesses of genotypes AA, Aa, and aa, respectively. The state transition functions of population genetics can hence be derived from the Price equation and the underlying causal model in the same fashion as in quantitative genetics.

5 Evolution as a Causal Process

The causal decompositions of the Price covariance given above reveal the causal structures underlying the evolutionary state transition functions and hence the evolutionary phenomena they describe. Our causal models satisfy all three desiderata mentioned earlier: they relate relevant genetic, phenotypic, and environmental factors; they give predictions of evolutionary consequences; and they can be used to estimate the effect of possible interventions on a subset of the variables. In addition to providing the causal foundations, the philosophical importance of defining the formal model is two-fold. First, it

tells us what selection must be in order for it to yield evolutionary change. Second, the explicit definition of the causal model makes it possible to determine whether fitness and/or selection cause evolution. These points are discussed in turn.

5.1 Selection as a causal process

All the causal models derived above required a trait to be a cause of fitness, favouring the notion of selection as a causal process (Millstein [2002], [2006]; Stephens [2004]), rather than a mere outcome (Matthen and Ariew [2009]; Matthen [2010]). The outcome interpretation claims that selection is nothing but a statistical fact holding in a population, such as the fitness variance or the fitness-trait covariance. At first sight such a view fits well with the popular accounts of selection, including Richard Lewontin's much-cited summary of Darwinian evolution as a necessary consequence of three conditions, phenotypic variation, differential fitness, and heritability, where differential fitness—that is, selection—means that 'different phenotypes have different rates of survival and reproduction in different environments' (Lewontin [1970], p. 1). In other words, phenotypes are correlated with fitness.

Our causal model, however, reveals an inadequacy of the purely statistical interpretation of adaptive evolution. To see this, imagine a situation where a trait does not cause fitness but both are affected by some common cause (Figure 5). Rausher ([1992]), for example, considers a hypothetical plant population whose foliar alkaloid concentration (phenotype) and seed production (fitness) are affected by the nitrate level of the soil environment (see also (Mauricio and Mojonniner [1997]; Morrissey *et al.* ([2010]) for similar discussions). The environmental confounder in such a situation will generate a statistical association between the trait and the fitness, so that Lewontin's criteria are satisfied, provided the trait is heritable. Evolutionary response,

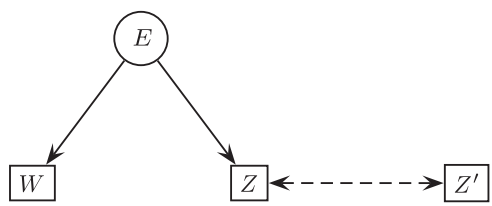

Figure 5. When the phenotype-fitness association is due only to a common cause, $\mathrm{Cov}(W, Z') = 0$ and no evolutionary response follows. But even in such cases, Lewontin's three conditions are satisfied and we may (falsely) assume a non-zero evolutionary response. Note that the path $W \leftarrow E \rightarrow Z \leftarrow \cdots \rightarrow Z'$ collides at Z and is not a trek. The dashed bidirected arrow represents reproductive pathways.

however, does not ensue for there is no trek between W and Z', and thus the Price covariance is zero. This simple example shows why the interpretation of selection as a pure outcome, as well as Lewontin's well-known formulation, is defective.[11] A mere statistical fact by itself has no explanatory role in the study of adaptive evolution.

The importance of distinguishing the selection-as-process from its statistical outcome cannot be emphasized too much. In a recent article, Sober ([2013]) correctly observed that fitness differences with respect to some trait Z_i do not entail a selection for Z_i itself, but wrongly concluded that they entail a selection for another trait Z_j that correlates with Z_i. It doesn't, since the phenotype–fitness association may be purely spurious, and in such cases there would be no evolutionary change. A mere statistical association between the trait and fitness itself does not imply any form of adaptive response; only selection-as-process does.

5.2 Causes of evolutionary change

Another contention in the statisticalist debate is whether fitness and/or selection can be regarded as a cause of evolutionary change (Stephens [2004]; Millstein [2006]; Otsuka *et al*. [2011]; Sober [2013]) or not (Matthen and Ariew [2002], [2009]; Walsh *et al*. [2002]; Walsh [2007], [2010]). The causal models provide a clear-cut solution to this entangled debate. Under the manipulationist account of causation, C is a cause of E if there is some intervention on C that alters the distribution of E (Woodward [2003]). In the same vein fitness is a cause of evolution if an intervention on W affects evolutionary response $\Delta\overline{Z}$. This can be easily verified by applying the standard intervention calculus (Pearl [2000]; Spirtes *et al*. [2000]) to the causal models defined above.

The post-intervention distribution can be represented by $P(\Delta\overline{Z}|do(W = w))$ where $do(\bullet)$ is Pearl's intervention operator that sets the fitness value to w. This amounts to forcing every individual in the population to have a certain number of offspring by some external means (for example, by culling all cubs after the wth birth). Alternatively, one can think of partial interventions that affect only some portion of the population. Assuming no individual gets more than one intervention, the result of partial interventions is given by the weighted average

$$P(\Delta\overline{Z}|\Omega) = \sum_{i}^{|\Omega|} \frac{n_i}{N} P(\Delta\overline{Z}|do(\omega_i)), \tag{20}$$

[11] Note that the case advanced here is to be distinguished from other criticisms of Lewontin's conditions, such as exact cancellation of selective force by other pathways (Wimsatt [1980a], [1980b]; Okasha [2007]), or an incidental trait–fitness correlation in a small population (Brandon [1990]). Lewontin's conditions may fail even in an infinite population undergoing no opposing evolutionary forces.

where $\Omega := \{do(\omega_1), do(\omega_2), \ldots\}$ is a set of partial interventions, N is the population size, and n_i is the number of individuals affected by $do(\omega_i)$. The standard intervention captured by Pearl's *do* calculus is just a special case of Equation (20) where Ω is a singleton. Here we consider only the standard ones. Our question thus amounts to whether $P(\Delta\overline{Z}|do(W = w)) \neq P(\Delta\overline{Z}|do(W = w'))$ for some $w \neq w'$.

So does an intervention on fitness affect evolution? It depends on the type of intervention. An intervention in a causal model is usually represented as a modification of the graph and/or the structural equations. 'Hard interventions' eliminate all the causal inputs to the target variables and impose a new set of values or distribution by some external force. In Figure 3, a hard intervention on fitness amounts to pruning all incoming arrows to W. This effectively interrupts all the treks from W to Z' so that the Price covariance becomes zero, that is, there is no evolutionary response. We thus conclude that hard interventions on W do not induce evolutionary change. This should not surprise us, for it is just a population-level restatement of Weismann's principle that no epigenetic surgery on parents would affect offspring phenotype. One can easily check that under the standard model this holds true for any phenotype, that is, $P(\Delta\overline{Z}|do(Z = z)) = P(\Delta\overline{Z}|do(Z = z'))$ for any hard intervention on Z.

From another perspective, however, this may appear puzzling: isn't artificial selection conducted by breeders a mixture of partial hard interventions? And we know that their efforts have considerably improved a number of phenotypes of economic importance, such as cows in terms of their milk yield. In these planned breedings, however, the intervention is a function of the phenotype: the breeder decides how many offspring an animal can have based on its phenotype. This effectively creates a new causal path from Z to W, that is, another selective pressure that leads to adaptive response. But unless the process itself is determined by the phenotype, hard interventions do not affect an evolutionary outcome.

Not all interventions are hard. 'Soft interventions' preserve some of the original causes of the target variable but modify its distribution, usually by adding another cause (Eberhardt [2007]). For example, suppose we want to know whether students' economic situation affects their academic performance. For this purpose, we may provide some financial aid or scholarship in order to see how this alters their exam scores. With respect to fitness, a soft intervention may be carried out through some form of environmental scaffolding (for example, additional food or provision of a nesting place), which is neither correlated with the focal phenotype nor interferes with its effect on the fitness. Such an independent additive intervention does not change the Price covariance, but does affect evolutionary responses through the mean fitness $\overline{W}$, the weighting factor in the Price Equation (8). If we boost fitness by

additive factor α, the post-intervention mean fitness becomes $\overline{W'} = \overline{W} + \alpha$, which results in a slower response to selection. In general, additive soft interventions on fitness conserve the direction but affect the rate of adaptive evolution.

Thus, there are some interventions on fitness that cause evolution. But it is important to note that not all interventions, even soft ones, induce population change. For example, if we manipulate only the variance of fitness by adding some noise factor with mean zero or by changing $\mathrm{Var}(E_W)$, these interventions will not affect either the Price covariance or the weighting factor $1/\overline{W}$. Hence, contrary to Sober's ([2013]) claim, the fitness variance does not cause evolutionary change, at least in case of directional selection.[12]

Finally, let us consider whether selection causes evolution. Selection, as discussed above, is a causal influence of the trait on fitness, whose linear magnitude is measured by coefficient β (Section 4.1). This parameter, in turn, should depend on selective environments including biotic (for example, prey abundance) as well as abiotic (for example, temperature) factors (Wade and Kalisz [1990]). Intervening on the selection-as-process thus amounts to a modification of these fitness-related factors controlling β. Since the Price covariance and the mean fitness are functions of β (Equation (12)), such interventions clearly make a difference in adaptive response. In general, we have

$$P(\Delta\overline{Z}|do(\beta)) \neq P(\Delta\overline{Z}|do(\beta')),$$

for any $\beta \neq \beta'$. It thus follows that selection does cause evolution.

To sum up: There are some interventions, either on fitness or on selection, that affect evolutionary response. Therefore, *pace* statisticalists, the causal model makes it clear that fitness and selection do cause evolution. But not every intervention will do. Hard interventions on fitness or manipulations of the fitness variance usually do not induce linear adaptive response. Let us emphasize that these conclusions were reached only with the aid of the causal models underlying the evolutionary formulae. A well-defined causal model gives an unequivocal answer to the question: 'does X cause Y?'. Purely conceptual analyses or interpretations of the putative cause, in contrast, never settle the issue.

6 Conclusion

In the history of evolutionary genetics, most of its celebrated principles have been formulated in probabilistic terms. The Price equation and Lewontin's

[12] If selection is acting on higher moments, as in stabilizing or disruptive selection, the fitness variance does matter to evolutionary change. Also, in a finite population the magnitude of drift is a function of the fitness variance. Sober's ([2013]) argument, however, focuses just on the linear adaptive response (that is, the breeder's equation).

conditions for evolution by natural selection both characterize evolution in terms of statistical, but not causal, features of a population. This gave rise to the philosophical puzzle as to whether evolution, described by these principles, is a causal process. The puzzle divided philosophers into two camps, but both sides seem to have accepted the statistical formulae as given and even admitted that the fundamental principles in evolutionary genetics are by nature non-causal or non-empirical. This presumption, however, is incorrect. As shown in this article, these evolutionary principles are derived from certain causal models, and in this sense not fundamental at all. What are really at the base of population change are the causal processes generating these statistics.

Like in many other cases, philosophers' *modus operandi* in this debate has been conceptual analysis. That is, the causal nature of selection or fitness was expected to be clarified by the correct interpretation of these concepts. To the eyes of these philosophers, the approach taken in this article may appear unfamiliar or even irrelevant. On the contrary, I argue that it is the only way to solve the issue: whether one variable causes another is answered not by interpreting these properties, but by specifying a causal model relating them. Once such a causal model is laid out, the answer follows quite straightforwardly.

In so arguing, I by no means pretend that the above models give the only causal structures underlying adaptive evolution; they are just a few—arguably the simplest—among many other possibilities. Nor am I trying to improve the predictive ability or performance of the standard evolutionary equations. My goal in this article was purely foundational, namely, to provide causal bases for the existing evolutionary formulae; no more, no less. A formal definition of the underlying causal model, however, proves useful in examining implicit assumptions and/or limitations of an evolutionary equation, for the graphical representation makes all the causal assumptions explicit. Figure 3 tells us, for example, that in order to apply the breeder's equation, the phenotype must cause fitness (a mere correlation is not sufficient), that its prediction eventually depends on the genotypic distribution (hence that the response may change across generations), and so on. The basic causal models constructed in this article also provide bases for more complex evolutionary phenomena, such as epigenetic inheritance or niche construction. These 'non-standard' mechanisms not covered by the traditional models introduce additional causal connections in the graph, whose impact on evolution can be directly evaluated through the method used in this article (Otsuka [2015]).

In sum, causal modelling provides a promising framework with which to approach a number of scientific, as well as philosophical, issues in evolution. Although its history dates back to Sewall Wright ([1921]), the technique has not received much attention either from biologists or philosophers until fairly

recently (for example, Shipley [2000], [2010]; Glymour [2006]). Exploring its possibilities and limitations will be important tasks for the future.

Acknowledgements

I owe a great debt to Lisa Lloyd for her extensive support during and beyond the writing process. Discussions with Bruce Glymour and Jim Griesemer were particularly helpful in forming and sharpening the ideas developed in this article. I also wish to thank the following people for their comments and discussions: Tyrus Fisher, Clark Glymour, Yoichi Ishida, Samuel Ketcham, Roberta Millstein, the members of the philosophy of biology lab at UC Davis, and the participants of the 2012 POBAM workshop at the University of Wisconsin-Madison. This work was partially supported by Japan Society for the Promotion of Science Grant-in-Aid for Postdoctoral Fellows (25-659).

Department of Philosophy
University of California, Davis
1240 Social Science and Humanities
One Shields Avenue, Davis, CA 95616, USA
junotk@gmail.com

References

Brandon, R. [1990]: *Adaptation and Environment*, Princeton: Princeton University Press.

De Jong, G. [1994]: 'The Fitness of Fitness Concepts and the Description of Natural Selection', *Quarterly Review of Biology*, **69**, pp. 3–29.

Eberhardt, F. [2007]: *Causation and Intervention*, Ph.D. dissertation, Carnegie Mellon University.

Fisher, R. A. [1930/2006]: *The Genetical Theory of Natural Selection*, Oxford: Oxford University Press.

Gillespie, J. H. [2004]: *Population Genetics: A Concise Guide*, Baltimore: Johns Hopkins University Press.

Glymour, B. [2006]: 'Wayward Modeling: Population Genetics and Natural Selection', *Philosophy of Science*, **73**, pp. 369–89.

Glymour, B. [2011]: 'Modeling Environments: Interactive Causation and Adaptations to Environmental Conditions', *Philosophy of Science*, **78**, pp. 448–71.

Lande, R. [1979]: 'Quantitative Genetic Analysis of Multivariate Evolution, Applied to Brain: Body Size Allometry', *Evolution*, **33**, pp. 402–16.

Lewens, T. [2010]: 'The Natures of Selection', *The British Journal for the Philosophy of Science*, **61**, pp. 313–33.

Lewontin, R. C. [1970]: 'The Units of Selection', *Annual Review of Ecology and Systematics*, **1**, pp. 1–18.

Lewontin, R. C. [1974]: *The Genetic Basis of Evolutionary Change*, New York: Columbia University Press.

Lloyd, E. A. [1988]: *The Structure and Confirmation of Evolutionary Theory*, Westport: Greenwood Press.

Lynch, M. and Walsh, B. [1998]: *Genetics and Analysis of Quantitative Traits*, Sunderland: Sinauer.

Matthen, M. [2010]: 'What is Drift? A Response to Millstein, Skipper, and Dietrich', *Philosophy and Theory in Biology*, **2**, p. e102.

Matthen, M. and Ariew, A. [2002]: 'Two Ways of Thinking about Fitness and Natural Selection', *The Journal of Philosophy*, **99**, pp. 55–83.

Matthen, M. and Ariew, A. [2005]: 'How to Understand Causal Relations in Natural Selection: Reply to Rosenberg and Bouchard', *Biology and Philosophy*, **20**, pp. 355–64.

Matthen, M. and Ariew, A. [2009]: 'Selection and Causation', *Philosophy of Science*, **76**, pp. 201–24.

Mauricio, R. and Mojonniner, L. [1997]: 'Reducing Bias in the Measurement of Selection', *Trends in Ecology and Evolution*, **12**, pp. 433–6.

Millstein, R. [2002]: 'Are Random Drift and Natural Selection Conceptually Distinct?', *Biology and Philosophy*, **17**, pp. 33–53.

Millstein, R. [2006]: 'Natural Selection as a Population-Level Causal Process', *The British Journal for the Philosophy of Science*, **57**, pp. 627–53.

Morrissey, M. B., Kruuk, L. E. and Wilson, A. [2010]: 'The Danger of Applying the Breeder's Equation in Observational Studies of Natural Populations', *Journal of Evolutionary Biology*, **23**, pp. 2277–88.

Okasha, S. [2007]: *Evolution and the Levels of Selection*, Oxford: Oxford University Press.

Otsuka, J. [2014]: *The Causal Structure of Evolutionary Theory*, Ph.D. dissertation, Indiana University.

Otsuka, J. [2015]: 'Using Causal Models to Integrate Proximate and Ultimate Causation', *Biology and Philosophy*, **30**, pp. 19–37.

Otsuka, J., Turner, T., Allen, C. and Lloyd, E. A. [2011]: 'Why the Causal View of Fitness Survives', *Philosophy of Science*, **78**, pp. 209–24.

Pearl, J. [1988]: *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, San Francisco: Morgan Kaufmann.

Pearl, J. [2000]: *Causality: Models, Reasoning, and Inference*, New York: Cambridge University Press.

Pigliucci, M. and Kaplan, J. [2006]: *Making Sense of Evolution: The Conceptual Foundations of Evolutionary Biology*, Chicago: University of Chicago Press.

Price, G. R. [1970]: 'Selection and Covariance', *Nature*, **227**, pp. 520–1.

Rausher, M. D. [1992]: 'The Measurement of Selection on Quantitative Traits: Biases due to Environmental Covariances between Traits and Fitness', *Evolution*, **46**, pp. 616–26.

Robertson, A. [1966]: 'A Mathematical Model of the Culling Process in Dairy Cattle', *Animal Production*, **8**, pp. 95–108.

Shipley, B. [2000]: *Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations and Causal Inference*, Cambridge: Cambridge University Press.

Shipley, B. [2010]: *From Plant Traits to Vegetation Structure: Chance and Selection in the Assembly of Ecological Communities*, Cambridge: Cambridge University Press.

Sober, E. [1984]: *The Nature of Selection: Evolutionary Theory in Philosophical Focus*, Chicago: University of Chicago Press.

Sober, E. [1993]: *Philosophy of Biology*, Boulder: Westview Press.

Sober, E. [2013]: 'Trait Fitness Is Not a Propensity, but Fitness Variation Is', *Studies in History and Philosophy of Biological and Biomedical Sciences*, **44**, pp. 336–41.

Spirtes, P., Glymour, C. and Scheines, R. [2000]: *Causation, Prediction, and Search*, Cambridge: MIT Press.

Stephens, C. [2004]: 'Selection, Drift, and the "Forces" of Evolution', *Philosophy of Science*, **71**, pp. 550–70.

Wade, M. J. and Kalisz, S. [1990]: 'The Causes of Natural Selection', *Evolution*, **44**, pp. 1947–55.

Walsh, D. M. [2007]: 'The Pomp of Superfluous Causes: The Interpretation of Evolutionary Theory', *Philosophy of Science*, **74**, pp. 281–303.

Walsh, D. M. [2010]: 'Not a Sure Thing: Fitness, Probability, and Causation', *Philosophy of Science*, **77**, pp. 147–71.

Walsh, D. M., Lewens, T. and Ariew, A. [2002]: 'The Trials of Life: Natural Selection and Random Drift', *Philosophy of Science*, **69**, pp. 452–73.

Wimsatt, W. [1980a]: 'Reductionistic Research Strategies and their Biases in the Units of Selection Controversy', in T. Nickles (*ed.*), *Scientific Discovery: Historical and Scientific Case Studies*, Dordrecht: Reidel, pp. 213–59.

Wimsatt, W. [1980b]: 'The Units of Selection and the Structure of the Multi-level Genome', *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, **2**, pp. 122–83.

Woodward, J. B. [2003]: *Making Things Happen: A Theory of Causal Explanation*, New York: Oxford University Press.

Wright, S. [1921]: 'Correlation and Causation', *Journal of Agricultural Research*, **20**, pp. 557–85.