The interpretation of selection coefficients

N. H. Barton[1] and M. R. Servedio[2]

1. Institute of Science and Technology
Am Campus 1
A-3400 Klosterneuburg
Austria.

2. University of North Carolina
Coker Hall
CB# 3280
Chapel Hill
NC 27599 U.S.A.

Running head: Interpreting selection coefficients

Telephone: 43-2243-9000-1071
Email: nick.barton@ist.ac.at

- **Keywords**

Direct selection, Indirect selection, kin selection, multilocus selection

Abstract

Evolutionary biologists have an array of powerful theoretical techniques that can accurately predict changes in the genetic composition of populations. Changes in gene frequencies and genetic associations between loci can be tracked as they respond to a wide variety of evolutionary forces. However, it is often less clear how to decompose these various forces into components that accurately reflect the underlying biology. Here, we present several issues that arise in the definition and interpretation of selection and selection coefficients, focussing on insights gained through the examination of selection coefficients in multilocus notation. Using this notation, we discuss how its flexibility - which allows different biological units to be identified as targets of selection - is reflected in the interpretation of the coefficients that the notation generates. In many situations, it can be difficult to agree on whether loci can be considered to be under "direct" versus "indirect" selection, or to quantify this selection. We present arguments for what the terms direct and indirect selection might best encompass, considering a range of issues, from viability and sexual selection to kin selection. We show how multilocus notation can discriminate between direct and indirect selection, and describe when it can do so.

Introduction

It is straightforward to use mathematical models to describe the deterministic evolution of a large population - we simply follow the frequencies of all the various gene combinations as they change through time. However, it is not at all easy to find a clear and meaningful interpretation of the consequent mass of coupled nonlinear equations, and thus there are heated debates over the interpretation of population genetic models (e.g. Sober, 1984, Nowak et al., 2010, Abbott et al., 2010). How do traits that do not directly affect fitness, such as mate preference or recombination rate, evolve? Can altruistic traits that reduce their bearer's fitness evolve? What, indeed, do we mean by fitness? Other questions center on the interpretation of selection (Sober, 1984). Does selection act on genes, on individuals, or on sets of interacting individuals? When does selection act on a gene or trait? How can one distinguish direct from indirect selection, and how should these types of selection be defined? In such examples, it has proved difficult and contentious to relate verbal arguments to mathematical models, and to understand the inner workings of the models. The difficulty is greater when the model consists of a simulation or numerical calculation, but even when the model has an explicit analytic representation, the meanings of the various terms are rarely obvious.

We examine several of the questions arising in the interpretation of selection via insights gained from a particular type of multilocus analysis. We focus specifically on what can be learned from the interpretation of the "selection" coefficients that appear in a methodology for multilocus population genetics (Barton and Turelli, 1991; Kirkpatrick et

al., 2002). This notation gives an exact analysis of selection and recombination, and allows a close "quasi- linkage equilibrium" approximation when recombination is faster than selection. It was developed in a series of papers, originally dealing with spatial clines (Barton, 1983, 1986) and the evolution of quantitative genetic variation (Barton and Turelli, 1987), but applied more broadly by Barton and Turelli (1991). Similar general multilocus models were developed independently by Bürger (1991) and Christiansen (1999). The method was set out very generally, allowing for sex - linkage and other forms of transmission, by Kirkpatrick, Johnson and Barton (2002), who suggested some changes to the notation introduced by Barton and Turelli (1991).

The key idea in this multilocus methodology is to represent genotype frequencies by their moments, and fitness as a polynomial function of genotype; with a finite number of alleles and loci, this leads to a closed set of equations, with a simple form that can be generated by symbolic computation. Moments are defined relative to a reference point that consists of a set of allele frequencies. Recursions describing evolution include emergent coefficients that multiply the moments of the system (Eq. 1). The basic method is generalizable, and allows considerable flexibility in the choice of reference point, in the way fitnesses are represented, and in the choice of which individuals are taken as the unit of selection. Thus, the choice of definition in a particular problem is crucial to our interpretation of the coefficients that define fitness. However, until now the interpretation of these coefficients, indeed whether in some cases they have any clear biological interpretation at all, has been unclear.

As we examine how the coefficients generated by multilocus analysis can be interpreted, we discuss insights that are applicable to several general issues in the interpretation of selection. We illustrate these issues with a variety of examples: sexual selection on haploids, diploid viability, epistasis, reinforcement, and kin selection. We stress that there is considerable freedom in multilocus analyses to choose the unit of selection to suit a particular biological problem, and there is also freedom to choose a reference point against which the associations are defined. This flexibility, however, means that it is not straightforward to interpret the various coefficients; indeed while in some cases very useful conclusions can be made from examining the coefficients, in other cases they remain uninterpretable. The lens of multilocus analysis raises some particularly interesting issues with regard to direct and indirect selection. We discuss the meaning of these terms, when these may be said to act on particular sets of loci, and how in some cases the coefficients defined by the notation of Barton and Turelli (1991) can be interpreted as corresponding to these types of selection. Examination of these issues through multilocus notation can provide a unique perspective that advances our understanding of the complexities of selection.

Multilocus population genetics

■ Notation

Here we give a brief summary of the multilocus methodology, highlighting the differences in the notation for selection between Barton and Turelli (1991), and Kirkpatrick, Johnson and Barton (2002) (denoted for brevity by BT and KJB respectively). We briefly illustrate the use of these notations with the example of sexual selection in haploids. More details are given in the Appendix. The allelic value of a gene is denoted by X; in this paper, we assume two alleles per locus, and take $X = 0$ or 1. There is no difficulty (in principle) in extension to any (finite) number of alleles. We assume a gene is always at some locus, i, in the genome. In the more general KJB notation, any particular gene will be in some "context" - for example, a gene in a diploid male, on the chromosome inherited from the female parent, could be denoted by $\mathbb{i} = i_{mf}$. Such "genes in context" are termed "positions", and denoted by double-struck font, so that $\mathbb{U}$ can represent an arbitary set of genes in context.

Here, we are primarily concerned with autosomal genes in diploids, or equivalently, mated pairs of haploids. In these specific cases, we can also use the condensed BT notation, in which a * denotes genes inherited from the paternal genome or the haploid male parent. Thus, in this notation we would denote a set of genes U expressed in females, V expressed in males by U, V^* (i.e., genes from mothers and fathers, respectively, in a diploid).

Genotype frequencies are represented by their moments, which are defined with respect to a reference point, $\wp_{\mathbb{i}}$ (see discussion below and Appendix). The deviation of the allelic value from the reference point is denoted by $\zeta_{\mathbb{i}} = X_{\mathbb{i}} - \wp_{\mathbb{i}}$, and the association between a set of genes $\mathbb{U}$ is defined as $D_{\mathbb{U}} \equiv \mathbb{E}[\zeta_{\mathbb{U}}] \equiv \mathbb{E}[\prod_{\mathbb{i}\in\mathbb{U}} \zeta_{\mathbb{i}}]$. If the reference point is set to be the allele frequency, and the genes $\mathbb{i}$, $\mathbb{j}$ are in the same genome, then $D_{\mathbb{i}\mathbb{j}}$ is the usual pairwise linkage disequilibrium.

■ Selection: KJB notation

We consider selection in very general terms, as a process in which the composition of a population changes in proportion to the fitnesses of its members. This includes viability and fecundity selection, as well as sexual selection, acting through mating success. As we argue below, this broad definition may also include assortative mating, which does not change the frequencies of alleles, but does change the contribution of pairs of alleles to the next generation; in this case, the population may be considered to consist of sets of genes or individuals.

We define the relative fitness of a genotype $\underline{X}$ simply as the ratio between its frequency in the population, $g(\underline{X})$, and its contribution to the next stage in the life cycle,

$g'(\underline{X})$, and write this as a polynomial function of genotype. KJB define relative fitness as:

$$\frac{W}{\overline{W}} = \frac{g'(\underline{X})}{g(\underline{X})} \equiv 1 + \sum_{\mathbb{U}\subseteq\Omega} a_{\mathbb{U}} (\zeta_{\mathbb{U}} - D_{\mathbb{U}}) \qquad (1)$$

(The term $\zeta_{\mathbb{U}} - D_{\mathbb{U}}$ ensures that $\mathbb{E}\left[W / \overline{W}\right] = 1$). This expresses the relation between the relative contribution of some group of genes, and their allelic value, $\underline{X}$. It is completely general, provided that we define enough coefficients $a_{\mathbb{U}}$ to specify the fitness of any genotype. With two alleles per locus, we need to specify a coefficient at least for every subset of the set of all positions, Ω. As we discuss below, in general, the $a_{\mathbb{U}}$ will depend on genotype frequencies, if only because of the normalisation $E[W / \overline{W}] = 1$.

The change in allele frequencies is given simply by:

$$\Delta p_{\mathbb{i}} = D'_{\mathbb{i}} = \sum_{\mathbb{U}\subseteq\Omega} a_{\mathbb{U}} D_{\mathbb{U}\mathbb{i}} \qquad (2)$$

(changes to the associations are defined similarly; see KJB). Here, we assume that the reference point equals the initial allele frequency $\wp_{\mathbb{i}} = p_{\mathbb{i}}$, so $D_{\mathbb{i}} = 0$. If we consider a pair of genomes, as with sexual selection amongst haploids, or viability selection on diploids, the change in allele frequencies depends on terms representing interactions between genomes, $\mathbb{U} = \{U, V^*\}$.

We illustrate this notation using the example of sexual selection in haploids, concentrating on the evolution of an allele for female preferences. Here we assume that there is a single preference locus $\mathbb{p}$, expressed only in females, which is one out of the total set of loci expressed across both males and females, $\mathbb{U}$. The change in preference frequency in females can be written as:

$$\Delta p_{\mathbb{p}} = \mathbb{E}\left[\zeta_{\mathbb{p}} \frac{W}{\overline{W}}\right] = \sum_{\mathbb{U}} a_{\mathbb{U}} D_{\mathbb{p}\mathbb{U}} \qquad (3)$$

where $\mathbb{p} = p_f$. Since $\mathbb{p}$ may be an element of the subsets $\mathbb{U}$, the expansion of the sum will include the term $a_{\mathbb{p}} D_{\mathbb{p}\mathbb{p}}$, terms which combine $\mathbb{p}$ with all other subsets of $\mathbb{U}$, and all other subsets of $\mathbb{U}$ that do not contain $\mathbb{p}$. If the reference point is set to the initial allele frequency such that $\wp_{\mathbb{i}} = p_{\mathbb{i}}$, then $D_{\mathbb{i}} = 0$, and each term in Eqn (3) will only be retained if the association D in that term are among multiple loci expressed in each represented sex (e.g., if $\mathbb{V}$ contains more than one locus expressed in males, and $\mathbb{p}$ is only expressed in females, $D_{\mathbb{p}\mathbb{p}} \neq 0$, $D_{\mathbb{p}\mathbb{p}\,\mathbb{V}} = D_{\mathbb{p}\mathbb{p}} D_{\mathbb{V}} \neq 0$, but $D_{\mathbb{p}\,\mathbb{V}} = D_{\mathbb{p}} D_{\mathbb{V}} = 0$ since $D_{\mathbb{p}} = 0$). Evolution at the preference locus in males would still be given by Eq. 3, except that $\mathbb{p} = p_m$. In this case, different terms would be retained in the expansion of Eqn (3), since $\mathbb{p}$ is now a member of the set of loci expressed in males. Thus, now $D_{\mathbb{p}\,\mathbb{V}} \neq 0$, since $\mathbb{p}$ and $\mathbb{V}$ are expressed in the

same sex.

■ Selection: BT notation

In contrast to the KJB notation, Barton and Turelli (1991) gave a definition that was specifically designed for selection on pairs of haploid genomes. They define separate coefficients for each sex, where the set U represents genes in females, and V represents genes in males. Using a ^ to denote the coefficients in the BT notation instead of the KJB notation, they define the fitness of pairs of loci in males and females (these could be mated pairs of haploids or genes from maternally and paternally inherited loci in diploids):

$$\frac{W}{\overline{W}} = \frac{g'(\underline{X})}{g(\underline{X})} \equiv 1 + \sum_U \hat{a}_{U,\emptyset}(\zeta_U - D_U) + \sum_V \hat{a}_{\emptyset,V}(\zeta_V^* - D_V^*) + \sum_{U,V} \hat{a}_{U,V}(\zeta_U - D_U)(\zeta_V^* - D_V^*). \quad (4)$$

Here, the first position in the subscript of the $\hat{a}$ coefficients denotes the set of genes in females and the second position denotes the set in males (an * is used to refer to males when the position is not clear from the subscript). The coefficients $a_{\mathbb{U}}$ from KJB are separated into coefficients expressing the effect on fitness of female genotype, $\hat{a}_{U,\emptyset}$, of male genotype, $\hat{a}_{\emptyset,V}$, and of the interaction between them, $\hat{a}_{U,V}$. Thus, provided that there are initially no associations between genes in males and females (i.e., $D_{U,V} = 0$), the marginal fitness of female and male genotypes depends only on $\hat{a}_{U,\emptyset}$ and $\hat{a}_{\emptyset,V}$, respectively. As we shall see below, changes in allele frequency within each sex depend only on these marginal coefficients.

There is a simple relation between the BT and KJB coefficients:

$$a_{U,V} = \hat{a}_{U,V} \quad \text{for } U,\ V \neq \emptyset$$
$$a_{U,\emptyset} = \hat{a}_{U,\emptyset} - \sum_{V \neq \emptyset} \hat{a}_{U,V} D_V^* \qquad a_{\emptyset,V} = \hat{a}_{\emptyset,V} - \sum_{U \neq \emptyset} \hat{a}_{U,V} D_U \quad (5)$$

As we discuss below, these differences between the notations have important implications for the interpretation of these coefficients.

As before, the change in allele frequencies can be found by assuming that the reference point equals the initial allele frequencies, yielding

$$\Delta p_i = \sum_U \tilde{a}_{U,\emptyset} D_{Ui}\ , \quad (6)$$

where $\tilde{a}_{U,\emptyset} = (\hat{a}_{U,\emptyset} + \hat{a}_{\emptyset,V})/2$. We see that allele frequencies in each sex change only as a result of the coefficients $\hat{a}_{U,\emptyset}$, $\hat{a}_{\emptyset,V}$ that represent selection within that sex, and not on interactions $\hat{a}_{U,V}$. This is not the case for the KJB $a_{\mathbb{U}}$.

Returning to the example above of sexual selection in haploids for illustration, evolution at the preference locus *p* can be represented in the BT notation by:

$$\Delta p_p = \sum_{U} \tilde{a}_{U,\emptyset}\, D_{pU} \qquad (7)$$

Since the coefficients $\tilde{a}$ are averaged over both sexes we do not distinguish between the sexes in the sets of loci expressed in the associations (the Ds) in the expansion of Eqn 7, as occurs in the KJB notation (this was accounted for during the averaging). Terms in the expansion of Eqn. (7) are thus retained provided that for $S \subseteq U$, $\tilde{a}_{S,\emptyset} \neq 0$.

As we will see from the additional examples below, we can choose the entities that are selected - single genes, pairs of genes in a diploid individual, and so on. Fitness may be influenced by many factors including differential survival, differential contributions to a mating pool, and genotype-dependent migration to different habitats. The interpretation of the coefficients a_U and $\hat{a}_{U,V}$ will depend on these choices. After selection, in this general sense, occurs, the sexual life cycle is followed by recombination, and a change in reference point, as detailed in the Appendix.

Examples

In the following examples, we demonstrate some of the properties of the coefficients in the multilocus notation, present some of the issues surrounding their interpretation, and argue that different notations are better suited for different examples. While so doing, we dissuss several issues regarding the general interpretation of selection, including what units selection can be said to act on, how this changes our quantification of selection, and how it might be determined whether direct or indirect selection is acting on specific genes or set of genes. We reserve $a_{\mathbb{U}}$ and $\hat{a}_{U,V}$ for the specific definitions of KJB or BT (eqs. 1, 4). When we define specific examples we will use letters other than "a" to represent the coefficients, to clarify discussion of how these coefficients relate to $a_{\mathbb{U}}$ and $\hat{a}_{U,V}$. The KJB notation, using the a coefficients, leads to more straightforward recursions, and was developed to handle a broader range of biological situations than the BT notation (including non-random mating among diploids, genomic imprinting, polyploidy, etc). However, as we show below, the BT notation, using the $\hat{a}$ coefficients, can, in appropriate cases, be more readily interpreted.

■ Defining fitnesses: choice of reference points and of sets of genes

Specific definitions of fitness in a particular biological problem will vary depending on several choices, including the set of genes under consideration and the reference point. (Recall that the latter is the point against which the genetic state of the

population is defined, and in multilocus notation defines the associations *D*). While in some cases the choice of both of these seems clear, in other cases there is some leeway that will consequently alter the multilocus coefficients. We also point out that the choice of both of these factors will almost always make the multilocus coefficients depend on genotype frequencies, even when the fitnesses themselves are not frequency-dependent.

Reference points can be chosen arbitrarily. It is usually most convenient to set the reference point equal to the current allele frequency in a population, and in many cases this gives an intuitive meaning to the selection coefficients (see section on direct and indirect selection below). One must still choose, however, across what set of genes the allele frequency will be taken (the "context" in KJB). For example, allele frequencies may differ between the sexes, or between the genomes derived from the mother or father, or between microhabitats; one must decide whether to take the allele frequency within a gene's context as the reference point ($p_i = \mathbb{E}[X_i]$), or instead choose some wider average. Indeed, one might choose a fixed arbitrary value: for example, Cheverud and Routman (1995) define a measure of "physiological" epistasis, which is independent of allele frequencies, by setting $\wp_i = \frac{1}{2}$.

That the choice of reference point generally renders multilocus coefficients frequency-dependent can be illustrated by a very simple example, involving two genes at the same locus. This example can be considered as viability selection on a diploid individual, or, equivalently, viability selection, and sexual selection on a pair of haploid individuals. (We refer to the two genes as "female" or "male", but in the diploid interpretation they are derived from the female and the male gamete). Under either interpretation, we can write fitness as:

$$W = 1 + b_{i,\emptyset}\,(X_i - \omega_i) + b_{\emptyset,i}\,(X_i^* - \omega_i^*) + b_{i,i}\,(X_i - \omega_i)\,(X_i^* - \omega_i^*) \qquad (8)$$

Here, we assume that genotypes have a *fixed* fitness, meaning that the fitness is not frequency dependent and does not change through time (to within a scaling factor that includes density-dependent regulation), and take the ω_i and ω_i^* as arbitrary, but fixed values. Now, the selection coefficients a_U can only be fixed if we keep these reference points fixed; if they are set equal to the ω_i, ω_i^*, then the a_U are identical to the b_U. In this case of a single locus, the KJB and BT notations are equivalent: $\mathtt{a}_U = \hat{\mathtt{a}}_U$.

In general, however, the coefficients $\mathtt{a}$ and $\hat{\mathtt{a}}$ do depend on allele frequency. We can choose the reference point to equal the current allele frequencies (p_i=$\wp_i$), but in general this will not equal the fixed ω_i used to define fitness in Eq. 8 . We can see this by letting $\Delta_i = p_i - \omega_i$ (recall that $\zeta_i = X_i - p_i$):

$$W = 1 + b_{i,\emptyset}\,(\zeta_i + \Delta_i) + b_{\emptyset,i}\,(\zeta_i^* + \Delta_i^*) + b_{i,i}\,(\zeta_i + \Delta_i)\,(\zeta_i^* + \Delta_i^*) \qquad (9)$$

Now, $\overline{W} = 1 + b_{i,\emptyset}\,\Delta_i + b_{\emptyset,i}\,\Delta_i^* + b_{i,i}\,\Delta_i\,\Delta_i^*$, and $\overline{W}\,a_{i,\emptyset} = b_{i,\emptyset} + b_{i,i}\,\Delta_i^*$, $\overline{W}\,a_{\emptyset,i} = b_{\emptyset,i} + b_{i,i}\,\Delta_i$, $\overline{W}\,a_{i,i} = b_{i,i}$. Thus, even in this simplest case, the coefficients $\mathtt{a}$ and $\hat{\mathtt{a}}$ must be frequency-dependent, both because they are normalized relative to the mean fitness, and because the marginal fitness of each allele depends on how often it finds itself associated with the other allele, via the coefficient $b_{i,i}$.

Fitnesses, as defined by the $b_{\mathbb{U}}$, might well also depend on genotype frequencies. In this case the $\mathtt{a}$, $\hat{\mathtt{a}}$ would depend on genotype frequencies for two reasons, both because of the changing reference point (as described above), and because of the intrinsic frequency-dependence of genotype fitnesses. For example, pure assortative mating can be modelled by assuming that a fraction (1-α) of pairs mate randomly, whilst a fraction α mate with the same type. For a single haploid locus, assuming equal frequencies in the two sexes, and setting $\wp_{\mathbb{i}} = p_{\mathbb{i}}$:

$$\begin{aligned} \mathrm{W} &= (1-\alpha) + \alpha\left(\frac{\mathrm{X_i}\,\mathrm{X_i^*}}{\mathrm{p_i}} + \frac{1}{\mathrm{q_i}}(1-\mathrm{X_i})(1-\mathrm{X_i^*})\right) \\ &= \\ &(1-\alpha) + \alpha\left(\frac{1}{\mathrm{p_i}}(\mathrm{p_i}+\zeta_\mathrm{i})(\mathrm{p_i}+\zeta_\mathrm{i}^*) + \frac{1}{\mathrm{q_i}}(\mathrm{q_i}-\zeta_\mathrm{i})(\mathrm{q_i}-\zeta_\mathrm{i}^*)\right) \qquad (10) \\ &= 1 + \frac{\alpha}{\mathrm{p_i}\,\mathrm{q_i}}\,\zeta_\mathrm{i}\,\zeta_\mathrm{i}^* \end{aligned}$$

This ensures that $a_{i,\emptyset} = a_{\emptyset,i} = 0$, so that allele frequencies do not change; the coefficient $a_{i,i}$ that represents assortment is necessarily frequency-dependent, becoming extremely large when either allele is rare.

As well as being free to choose an arbitrary baseline, $\omega_{\mathbb{i}}$, and reference point, $\wp_{\mathbb{i}}$ for defining fitnesses, we can also choose which sets of genes to follow, both within and across loci. This will also affect definitions of fitness, and affect the multilocus coefficients. Suppose that instead of following the contribution of pairs of genes to the next generation, we follow just genes in females. Because we are now only following genes in one position (the females), we write coefficients as a_i, rather than $a_{i,\emptyset}$. From Eq. 9, their marginal relative fitness is now $W/\overline{W} = 1 + \left(b_{i,\emptyset} + b_{i,i}\,\Delta_i^*\right)\zeta_i/\overline{W}$, so that again $a_i = \left(b_{i,\emptyset} + b_{i,i}\,\Delta_i^*\right)/\overline{W}$: if there is an interaction between the genes, $b_{i,i}$, due to non-random mating between haploids or dominance in diploids, then the selection coefficient for genes in females depends on the frequency of the allele that they encounter in males $\Delta_i^* = p_i^* - \omega_i^*$. The change in allele frequency is $\Delta p_i = a_i\,D_{ii} = a_i\,p_i\,q_i$. If allele frequencies are the same in the two sexes, then this gives a closed equation for the allele frequencies, with the frequency-dependence that arises from dominance or non-random

mating between haploids being absorbed into the single coefficient $a_{i,\emptyset} = a_{\emptyset,i} = a_i$. If we have in mind viability selection on diploids, then it is natural to assume this symmetry: it would only be violated if genes were imprinted, so that they have different effects when inherited from the mother *versus* the father. However, if selection acts on pairs of haploid individuals, with distinct sexes (as in the original discussion of Eqn 9), then selection may well act differently on genes at the same loci that find themselves in a different sex.

Just as we must make a choice of what set of genes (in what context) to use as a reference point at a given set of loci, a choice must also be made of what loci to track. This choice is more critical. If we do not follow all of the genes that cause fitness differences, then the recursions across generations will still be correct, but the coefficients may change with the frequencies of genes or gene combinations that are not themselves tracked. A complete description of the dynamics requires that we follow *all* of the genes that affect fitness.

■ Direct vs. indirect selection

Evolution at a locus does not require that selection acts specifically on alleles at that locus, but can occur instead because of selection on other loci in the system. In other words, alleles that do not cause differences in fitness may change in frequency if they are correlated with fitness. In order to begin to partition out the pathways leading to evolution at a locus from different sources of selection, it is useful to distinguish between two different types of selection, direct and indirect. In population genetics, selection is said to act directly on a set of genes when those genes cause differences in fitness of the individual (or set of individuals) that carries them. Direct selection can lead to changes in allele frequency and/or changes in disequilibria; the latter can occur even when allele frequencies themselves do not change. Selection acts indirectly on genes that do not affect the fitness of their carrier; linkage disequilibrium (or more generally, associations) between the focal loci and other genes under selection is required for indirect selection to operate. The distinction between direct and indirect selection has long been critical in the field of sexual selection, where they distinguish two biologically independent modes of evolution of female preferences: direct selection on preferences occurs when specific preference alleles have effects on female survival or fecundity (Kirkpatrick and Ryan 1991). In contrast, indirect selection on a preference is defined as occuring when preferences change in frequency due to the presence of genetic associations between the preference and other loci, such as male display loci, that are under selection.

We note that the distinction between direct and indirect selection tends to be made differently in the literature on social evolution. The terms in Hamilton's Rule correspond to the direct effects of genes on the individual that carries them, plus the indirect effects on fitness of genes in other individuals; these genes are typically at the same locus. In

population genetics, indirect selection refers to the effects of genes at different loci, but in the same individual.

Modifier loci provide another interesting example in which to consider direct and indirect selection. Modifiers have sometimes been defined to be neutral with regard to fitness, which implies that they can only be under indirect selection. However, they can be defined more generally, as alleles which alter some aspect of the genetic system (Feldman et al. 1996). Under this latter definition, selection on a modifier can be direct, because a modifier can alter the fitness of an individual that carries it. Modifiers of dominance and epistasis are necessarily directly selected, because they alter the fitness of an individual in a way that depends on other genes, at the same or at different loci. We discuss modifiers of dominance (more precisely, of heterozygote fitness) below. Other examples in which selection can be direct include modifiers of ploidy, and modifiers that affect migration.

Here, we illustrate that when the reference point is chosen carefully, it is possible for the coefficients to provide measures of direct and indirect selection on genes; these are composite measures of selection with a clear biological interpretation. Once we have chosen the reference point, and the sets of genes that we follow, the coefficients are uniquely defined. We show that direct and indirect selection can, in some circumstances, be distinguished by these coefficients using the BT formulation of the multilocus notation.

First, we repeat a word of caution that extends to all examples below - unless we have included *all* the genes that cause fitness differences in the set that we follow ("causal loci"), we will not be able to determine which are directly versus indirectly selected, and we will not have a closed set of equations that allow us to predict evolution over multiple generations. This can easily be seen if we think of a set of neutral markers, that are associated with unobserved selected loci - the typical situation with real data. We can define, in any one generation, the average fitness of each marker genotype, and hence the a's, but have no way to predict how these coefficients will change from one generation to the next, since the markers do not cause the fitness differences. Below, we use examples to illustrate several other general issues that arise in attempting to distinguish direct and indirect selection using multilocus notation.

- **Sexual selection in haploids**

If all causal sets of genes in haploids are included in our definition of fitness, then we can identify which sets of genes are directly selected - that is, which genes influence the fitness of the individuals that carry them. Here, we use the model of sexual selection in haploids introduced above (see Notation), where we are interested in evolution at a locus for female preference, which together with its context is denoted $\mathbb{p}$ in the KJB and p in the BT notation. We assume that the preference locus itself is not under viability or fecundity selection, and that this is the only locus expressed in females. We also assume that all

causal loci in the system are included in p in females and the set V in males (e.g., p is the only member of set U in Fig. 1a). Assuming strict polygyny, where all females have equal mating success, there is thus no direct selection on females. We will show that this is reflected in the multilocus coefficients from the BT notation. The fitness of haploid pairs must have the form:

$$\frac{W}{\overline{W}} = 1 + \sum_V \hat{a}_{\emptyset,V}(\zeta_V^* - D_V) + \sum_{p,V} \hat{a}_{p,V}(\zeta_p - D_p)(\zeta_V^* - D_V) \quad (11)$$

That is, in the BT notation, the coefficient $\hat{a}_{p,\emptyset} = 0$, reflecting the lack of direct selection at the p locus in females (recall that gene for female preference, p, is not expressed in males in this example). Because $\hat{a}_{p,\emptyset} = 0$ in this example, the frequency of the locus p changes only in so far as it is associated with other directly selected sets. In other words, any evolution that occurs at the preference locus from Eq. 7, using the BT notation, must come through associations of p with selected loci, V (e.g., through D_{pV}); it is thus indirect.

In contrast, the lack of direct selection on locus p is less obvious in the more general KJB notation. Now, we define relative fitness as in Eqn 1, where $\mathbb{U}$ includes genes in both males and females. Because $a_{U,\emptyset} = \hat{a}_{U,\emptyset} - \sum_U \hat{a}_{U,V} D_V^*$, when $D_V^* \neq 0$, $a_{U,\emptyset}$ is not necessarily zero, so the fact that preference alleles themselves do not affect fitness of females is not reflected in the $a_{U,\emptyset}$. In fact, in our sexual selection example, $a_{p,\emptyset}$ will necessary not equal zero if females prefer a combination of genes in males V, since both $\hat{a}_{p,V}$ and D_V^* will be non-zero.

The change in preference frequency in females with the KJB notation can be seen in Eq. 3 above. This expression is very simple, but it is not as obvious as with the BT notation that when D_p, $D_{p,V} = 0$, then $\Delta p_{p,\emptyset} = 0$. To see this, substitute $a_{U,V} = \hat{a}_{U,V}$, $a_{\emptyset,V} = \hat{a}_{\emptyset,V} - \sum_U \hat{a}_{U,V} D_U$, and $a_{U,\emptyset} = \hat{a}_{U,\emptyset} - \sum_U \hat{a}_{U,V} D_V^*$ into Eq. 3, giving:

$$\Delta p_{p,\emptyset} = \sum_U \hat{a}_{U,V} D_{pU,V} + \sum_V \left(\hat{a}_{\emptyset,V} - \sum_U \hat{a}_{U,V} D_U\right) D_{p,V} + \sum_U \left(-\sum_U \hat{a}_{U,V} D_V^*\right) D_{pU} \quad (12)$$

Since $D_{p,V} = 0$ and $D_{pU,V} = D_{pU} D_V$, this is zero, as required.

The example above shows that even when the BT coefficient $\hat{a}_{U,\emptyset} = 0$, the KJB coefficient may be non-zero, so that direct selection is not reflected in the KJB notation.

We see that for haploids, the most natural representation is that of Barton and Turelli (1991). Note, however, that even with the BT notation, $\hat{a}_{U,\emptyset}$ only accurately represents direct selection on the set U in haploid females provided that, as specified above, all causal loci influencing fitness in females are included in the set U.

Indirect selection, like direct selection, can again be represented by the BT notation. In the case of selection on locus i, the term $\sum_U \hat{a}_{U,\emptyset}\, D_{Ui}$ describes the evolution at locus i that occurs because of selection on the set U, transmitted to i by the genetic association between i and U. Because the selection coefficient $\hat{a}_{U,\emptyset}$ accurately represents direct selection on the set of genes U in this notation, it also represents the indirect selection that the set U will transmit to gene i. The measures of selection, both direct and indirect, extracted by the coefficients are an agglomeration of frequency-dependent effects on fitness from viability selection, sexual selection, and fecundity selection, if it is acting. The coefficients thus allow the total evolutionary change to be separated into components that can be attributed to the loci through which the selection originates.

- **Epistasis between three loci**

An attempt to specify whether particular multilocus coefficients impart direct or indirect selection on a locus quickly brings to light an interesting issue in defining whether sets of loci are under direct or indirect selection. Provided that we have included all causal loci, we can identify which complete sets of genes are directly selected. However, one must be very careful with language in stating whether particular *subsets* of a set of genes are also directly selected.

Consider an example with selection on sets of genes, where we do not need to consider separate sexes. In this simplest case, there is no distinction between the KJB and BT notations. In such an example, if $a_i = 0$, we might say that gene i is not directly selected. However, the entire set of genes $\{i,\ j,\ k\}$ may be directly selected, meaning that certain combinations of alleles will have high fitness (e.g. if $a_{ijk} \neq 0$). In this case, allele i will change in frequency even if $a_i = 0$, and even if it is in linkage equilibrium with both j and k ($D_{ik} = D_{ij} = D_{ijk} = 0$) - provided that j and k are associated with each other: $\Delta p_i = \tilde{a}_{ijk} D_{iijk} = \tilde{a}_{ijk}(p_i q_i D_{jk} + (q_i - p_i) D_{ijk})$. This can easily be seen from Table 1, which presents the a's for this problem: amongst genotypes X_j, $X_k = 0, 0$ or $1,1$, allele $X_i = 1$ always increases fitness, and so if $D_{jk} > 0$, it will increase. Since the change in frequency at locus i occurs because of direct selection on a set of which i is a subset, we can say that i ultimately *evolves because of* direct selection. We cannot say, however, that there is direct selection *specifically on the locus i*. Note that since the change in frequency at locus i arises in the absence of a genetic association between i and another locus we cannot define the selection on i as indirect.

Table 1. Relative fitnesses in a three locus example with epistasis, where $W / \overline{W} = 1 + \hat{a}_{ijk}(\zeta_{ijk} - D_{ijk})$

X_i	X_j	X_k	$W / \overline{W}$
0	0	0	$1 - a_{ijk}\, p_i\, p_j\, p_k - a_{ijk}\, D_{ijk}$
1	0	0	$1 + a_{ijk}\, q_i\, p_j\, p_k - a_{ijk}\, D_{ijk}$
0	0	1	$1 + a_{ijk}\, p_i\, p_j\, q_k - a_{ijk}\, D_{ijk}$
1	0	1	$1 - a_{ijk}\, q_i\, p_j\, q_k - a_{ijk}\, D_{ijk}$
0	1	0	$1 + a_{ijk}\, p_i\, q_j\, p_k - a_{ijk}\, D_{ijk}$
1	1	0	$1 - a_{ijk}\, q_i\, q_j\, p_k - a_{ijk}\, D_{ijk}$
0	1	1	$1 - a_{ijk}\, p_i\, q_j\, q_k - a_{ijk}\, D_{ijk}$
1	1	1	$1 + a_{ijk}\, q_i\, q_j\, q_k - a_{ijk}\, D_{ijk}$

More specifically, this issue will arise whenever we try to identify whether there is direct selection on a particular *subset* of loci, when the set as a whole influences fitness. Consider four loci, i, j, k, and l. If $a_{ijkl} = 0$ and $a_k = 0$, but $a_{ijl} \neq 0$, we can say that there is no direct selection on locus k, and that there is direct selection on the set $\{i, j, l\}$, but will not be able to say that specific subsets of $\{i, j, l\}$ are directly selected if the as that correspond to those subsets are zero. We can only say that they evolve because of direct selection on a larger, directly selected set. Note that if the a that corresponds to a subset of $\{i, j, l\}$ is non-zero, we can say that that a represents a *component* of direct selection on that subset.

Hastings (1985) and Barton (1986) showed that when selection is weak relative to recombination, selection a_U for a set of genes U in haploids generates linkage disequilibrium, D_U, amongst precisely those genes (to leading order in the strength of selection). In this limiting case, we can interpret the coefficients a_U (or $\hat{a}_U$) as directly selecting for the corresponding D_U. This argument can be extended to allow strong directional selection (i.e. $a_i >> a_U$), but then, the simple relation between the epistatic coefficients and the associations that they produce is lost (see Appendix).

■ Multiple interpretations of the same coefficients: reinforcement

In some situations, examination of the multilocus notation can point out parallels between very different biological processes. The biological interpretations of the multilocus notation must thus be made in context, because the same coefficients may represent different natural processes. We illustrate this with a simple model of the components of the process of reinforcement. The coefficient $a_{i,i}$ may represent either assortative mating between haploid individuals, or reduced survival of the heterozygotes formed in diploids after meiosis. In the classical view of reinforcement, we expect that

selection can only increase prezygotic isolation, but cannot decrease the survival of heterozygotes (Dobzhansky, 1940). Yet, both processes are represented by an increase in $a_{i,i}$. The difference lies in the constraints that we implicitly assume. It is natural to suppose that an increase in assortative mating increases the contribution of homozygous pairs to the next generation, and decreases the contribution of heterozygous pairs, without changing allele frequencies. In contrast, we could hardly assume that a decrease in survival of heterozygotes would be compensated by an increase in survival of homozygotes: the latter we assume to be held fixed at some maximum value.

To illustrate how the same coefficients take on different meanings due to constraints, consider the fitness of an additive modifier k of assortment between haploids at locus i would be written as

$$\frac{W}{\overline{W}} = 1 + \frac{\alpha_0(2-X_k-X_k^*)+\alpha_1(X_k+X_k^*)}{2\,p_i\,q_i}(X_i - p_i)\,(X_i^* - p_i) \tag{13}$$

which implies BT coefficients $\hat{a}_{i,i} = (\alpha_0\, q_k + \alpha_1\, p_k)\,/\,(p_i\, q_i)$,

$\hat{a}_{ik,i} = \hat{a}_{i,ik} = (\alpha_1 - \alpha_0)\,/\,(2\, p_i\, q_i)$, assuming reference points equal to allele frequencies, and allele frequencies the same in each sex. This definition ensures that marginal fitnesses at locus i do not vary, so that there is pure assortment, with no sexual selection. This can be interpreted as a fraction α of individuals mating assortatively, and the remainder randomly (O'Donald, 1980). There is no direct selection on the modifier $\left(\hat{a}_{\emptyset,k} = \hat{a}_{k,\emptyset} = 0\right)$.

By contrast, an additive modifier k of heterozygote fitness would be written as:

$$\begin{aligned}\frac{W}{\overline{W}} &= 1 - (s_0(2 - X_k - X_k^*) \\ &\qquad + s_1(X_k + X_k^*))\,(X_i(1 - X_i^*) + (1 - X_i)\,X_i^*) \\ &= 1 \\ &-\overline{s}((\zeta_i + \zeta_i^*)\,(q_i - p_i) - 2\,\zeta_i\,\zeta_i^*) - \\ &\quad (s_1 - s_0)\,(\zeta_k + \zeta_k^*)\,(2\,p\,q + (\zeta_i + \zeta_i^*)\,(q_i - p_i) - 2\,\zeta_i\,\zeta_i^*)\end{aligned} \tag{14}$$

where $\overline{s} = s_0\, q_k + s_1\, p_k$, $\overline{W} = 1 - 2\,\overline{s}\,p_i\,q_i$

so that the selection against heterozygotes $(s_0(2 - X_k - X_k^*) + s_1(X_k + X_k^*))$ is $2\,s_0,\ s_0 + s_1,\ 2\,s_1$, depending on whether the modifier locus has value $X_k + X_k^*$=0, 1 or 2. This implies coefficients $\overline{W}\,\hat{a}_{i,\emptyset} = \overline{W}\,\hat{a}_{\emptyset,i} = 2\,\overline{s}(p_i - q_i)$, $\overline{W}\,\hat{a}_{i,i} = 2\,\overline{s}$,

$\overline{W}\,\hat{a}_{\emptyset,k} = \overline{W}\,\hat{a}_{\emptyset,k} = -2\,pq(s_1 - s_0)$, and $\overline{\mathtt{W}}\,\hat{\mathtt{a}}_{\mathtt{i,k}} = \overline{\mathtt{W}}\,\hat{\mathtt{a}}_{\emptyset,\mathtt{ik}} = \overline{W}\,\hat{a}_{k,i} = \overline{W}\,\hat{a}_{ik,\emptyset} =$ $(s_1 - s_0)\,(p_i - q_i)$, $\overline{W}\,\hat{a}_{i,ik} = \overline{W}\,\hat{a}_{ik,i} = 2\,(s_1 - s_0)$. Thus, selection now acts directly on the modifier, via the coefficients $\hat{a}_{\emptyset,k},\ \hat{a}_{k,\emptyset}$. (For a full model of reinforcement, we would assume that locus i is associated with other genes, U, under selection. Then, isolation may be strengthened either by a modifier of assortment, or by an increase in associations, D_{iV} for $V \subseteq U$. These correspond to Felsenstein's (1981) one-and two allele models,

respectively; Barton and De Cara, 2010). The key point here is that the appropriate definition of the coefficients depends on the constraints placed on them, which depends in turn on the biological context.

■ Units of selection: Kin selection

It is well understood that selection may act at different levels or on different units. There are interesting parallels both in the issues that arise when considering selection on different units and in the ways that this selection is defined notationally. Kin selection, for example, is a kind of indirect selection, in which an allele spreads by assisting the reproduction of other individuals that tend to carry the same allele. It is commonly represented by Hamilton's Rule (Hamilton, 1964a,b), which states that a trait will increase if its direct cost is outweighed by the benefit it gives to others, multiplied by their "relatedness". This term is misleading, because "relatedness" need not be due to simple kinship, and the coefficient of relatedness in general depends on the whole evolutionary process, not just on the pedigree: thus, it will vary across the genome, and can only be estimated from the pedigree when selection is weak (Queller, 1992; Gardner et al., 2011). There are further difficulties in understanding the "inclusive fitness" of an individual, which attempts to identify the causal effect of that individual on the propagation of the alleles that it carries, through both its own and the others' reproduction. These difficulties have led to heated debate; for a recent manifestation, see Nowak et al. (2010) and Abbott et al. (2010).

Kin selection can be represented using the multilocus formalism, with the associations $D_{\mathbb{U}}$ playing the role of "relatedness" (Gardner et al., 2007). In general, the fitness of each reproductive individual depends on both its own genotype, and on the genotypes of all its neighbours. However, we can illustrate the issues by considering just pairwise interactions. The fitness of a reproductive individual depends on its own genes, X, and on the genes of another randomly chosen 'helper', X^*. We then show how (just as with sexual selection between a pair of haploid individuals) direct selection can be accurately represented by an extension of the BT notation. There may also be recombination, union of gametes, mixing of subpopulations, or further rounds of selection: we just focus on one round of kin selection that involves interactions between random pairs.

We define the fitness of a set of individuals as their relative contribution to the next stage of the life cycle. Using the general KJB formalism, fitness is defined by coefficients $a_{\mathbb{V}}$ (Eq. 1), with the set $\mathbb{V} = \{U, V\}$ of genes including a set U from the reproductive individual, and V from the helpers (Fig. 1b). Alternatively, fitnesses could be defined in the BT notation, using coefficients $\hat{a}_{U,V}$ (Eq. 4). Clearly, the genes in the two sets U, V need not be at the same loci.

Assume for the moment that genomes $\underline{X}$, $\underline{X}^*$ are associated randomly, so that initially $D_{U,V} = D_U D_V$. Then, a gene i can only increase *via* the reproductive individual, at a rate $\Delta p_i = \sum_U \hat{a}_{U,\emptyset} D_{iU}$ (Eq. 6), which is independent of both $\hat{a}_{U,V}$ and $\hat{a}_{\emptyset,V}$. Moreover, associations amongst genes in the reproductive individual (which are the only associations that can be transmitted to future generations) change as $\Delta D_S = \sum_U \hat{a}_{U,\emptyset}(D_{SU} - D_S D_U)$ (Kirkpatrick et al., 2002, Eq. 9), which again does not depend on $\hat{a}_{U,V}$ or $\hat{a}_{\emptyset,V}$. This is simply because the output from the reproductive individuals depends only on the $\hat{a}_{U,\emptyset}$, and is only influenced by the 'helpers' in so far as these affect those $\hat{a}_{U,\emptyset}$. The coefficient $\hat{a}_{U,\emptyset}$ is thus akin to a measure of direct selection; direct selection as acting only when genes in the reproductive individual directly affect that individual's fitness.

Kin selection requires that there be associations between genes in individuals that interact (i.e., "relatedness"), so that immediately before kin selection we have $D_{U,V} \neq 0$. Then, $\Delta p_i = \sum_U \hat{a}_{U,\emptyset} D_{iU} + \sum_V \hat{a}_{\emptyset,V} D_{i,V} + \sum_{U,V} \hat{a}_{U,V}(D_{iU,V} - D_{iU} D_V - D_U D_{i,V})$, showing that an allele at locus i can increase through indirect selection if it is associated with genes in helpers ($D_{i,V}$, $D_{iU,V}$).

Effects of the helpers on the fitness of the reproductive individual can thus appear in two ways. First, they make the marginal fitnesses of alleles in the reproductive individual depend on allele frequencies in the helpers (making the coefficients of direct selection, $\hat{a}_{U,\emptyset}$, frequency-dependent). Second, if there are associations between genes in the helpers and allele i in the reproductive individual, ($D_{i,V}$, $D_{iU,V}$), then they cause indirect selection on allele i. Just as with sexual selection between haploids, this definition corresponds to the BT notation, but not the KJB notation.

This formalism might represent classical kin selection, but could equally well be thought of as describing frequency-dependent selection. Less obviously, it could represent the effect of genes in other species on the focal allele. The formalism is agnostic as to the causes of the associations $D_{U,V}$, which might be due to kinship, to selection or assortment at some previous stage, or to some kind of habitat choice. A closed model requires that we explain the evolution of these associations (Gardner et al., 2007). With weak selection, quasi-linkage equilibrium can be used to approximate the "relatedness" due to kinship, in the strict sense of pedigree relationship; associations due to habitat choice within or between species are generated directly by genotypes that affect that choice.

■ Alternate notation: Viability and sexual selection on diploids

In some cases, neither the BT nor the KJB notation suits the needs of the problem, and we must turn to an alternate notation. Another representation is most suitable, for

example, when we want coefficients to specifically represent additive and dominance effects during viability selection in diploids. If there is no imprinting, and no difference between *cis* and *trans* genotypes, then fitness depends only on the diploid genotype at each locus, and can be written as:

$$W = 1 + \sum_{A,B} \alpha_{A,B} \prod_{i \in A} (\zeta_i + \zeta_i^*) \prod_{i \in B} (\zeta_i \, \zeta_i^*) \qquad (15)$$

Here, the coefficient $\alpha_{A,B}$ represents an interaction between the additive effects of the set of loci A, and the dominance effects of loci in the set B; A, B include the null sets but otherwise do not overlap (Fig. 1c). If all $\alpha_{A,B} = 0$ for $B \neq \emptyset$, there is no dominance, but there can be arbitrary epistasis between the additive effects of the loci. However, these coefficients depend on the choice of reference points, and so the partitioning will change with allele frequencies (see Barton and Turelli, 2006).

It may be appropriate to define such specialised notations for specific problems, but one can always write the more general $a_{\mathbb{U}}$ in terms of such coefficients, and then use the general KJB machinery to find how the population evolves. In this example, the coefficients in either of the standard notations (i.e. $a_{\mathbb{U}}$, $\hat{a}_{U,V}$) have a complex dependence on the $\alpha_{A,B}$, and do not reflect the additive and dominance components of selection in a straightforward way. Since the general recursions for the KJB notation are simpler than those for BT, it is simpler to use the former in calculations.

If sexual selection between diploids is considered we can again use Eq. 4, but the sets U and V have a different interpretation. Each set now represents not only the loci within a particular sex, but the pairs of maternally and paternally inherited positions within each locus (as in the KJB notation). Provided that the $\hat{a}$ are calculated as in the example of sexual selection in haploids above, coefficients corresponding to direct and indirect selection can still be identified in this extension of the BT notation, although these types of selection can have several components, for which it may not be possible to determine specific interpretations.

Summary

Descriptions of selection can sometimes seem straightforward; the relative fitness of two genotypes at a locus from a specific bout of selection, for example, can be described by a simple ratio. The interpretation of selection can, however, be considerably more complicated, particularly when multiple loci are considered. We use multilocus notation, in which evolutionary changes in gene frequencies and genetic associations are written in terms of the products of frequency-dependent selection coefficients and moments of gene frequencies, to discuss several general issues surrounding the

interpretation of selection. First, we point out that the choices made in the set of genes considered and the set of allele frequencies that constitute a reference point can alter descriptions of selection, and consequently the interpretation of the multilocus selection coefficients. We show, however, that in some cases these choices can be made in a way that yields selection coefficients with intuitive meanings.

Another major focus of our discussion is the identification of sources of direct and indirect selection. The definition of direct selection, when focussed on a set of genes in one sex, seems clear; selection on a set of genes in an individual is direct when that set of genes affects the fitness of that individual, in the absence of genetic associations between that set of genes and other sets of genes. We show above, however, that one must be very cautious in describing whether particular subsets of loci are under direct selection. In contrast to direct selection, indirect selection can be thought of as having three distinct meanings, all of which are mediated by associations between sets of genes. First, indirect selection can act on genes that do not themselves cause any changes in fitness, but are correlated with selected genes as a result of their own effects (e.g., selectively neutral female preference, Kirkpatrick and Ryan 1991). Second, it can act on genes that do not cause any changes in fitness, but are correlated by chance with selected genes (e.g. genes caught in selective sweeps). Finally, in social evolution, genes in an individual are said to be indirectly selected when they cause changes in the fitness of another individual, with whom the focal individual shares genes. Indirect selection thus seems equivalent to the current usage of the term kin selection, in which "relatedness" is a measure of the associations between sets of genes in different individuals (Gardner et al., 2007, 2011).

We demonstrate that Barton and Turelli's (1991) formulation of the multilocus notation can accurately partition out selection coefficients for direct and indirect selection on inclusive sets of causal loci in all three circumstances, given a sensible choice of reference point, due to the fact that it is formulated to describe mated pairs of haploids. This can lead to important insights into the sources of selection on loci in haploid models. In the case of the evolution of male mating preferences, for example, the BT notation can separate out direct selection against preference alleles, due to increased competition for mates, from indirect selection transmitted to the preference locus via genetic associations with a male trait; it thus allows a comparison of the magnitude of these effects (Servedio 2007). Although formulated for mated pairs of haploids, we believe that by extension, it is possible to modify the same notation to describe pairs of diploids, or indeed, of arbitrary sets of genes in a way that retains the meaning of coefficients with regard to direct and indirect selection: fitness should be defined in terms of products across classes of gene, as in Eq. 4).

The fact that selection coefficients describing direct and indirect selection during social evolution can be identified illustrates parallels in the interpretation of selection

across different units of selection. Such parallels should also be able to be drawn at other levels of selection, again given a meaningful choice of reference point; in these cases the selection coefficients should again have a ready interpretation using the BT notation. In other cases, regardless of the unit of selection, the more flexible KJB notation may be simpler. It may then, however, not be possible to interpret selection coefficients in terms of direct and indirect selection. In yet further cases, such as the example of additivity and dominance in diploids above, customized notation may be developed that can have a meaningful interpretation for a given problem.

Many of the issues we have raised parallel those discussed in the philosophical literature. Our perspective is essentially the same as that of Sober and Lewontin (1982). It is always possible to assign selection coefficients to any set of alleles, such that these determine the change in frequency of these alleles. However, these coefficients will change over time, and do not necessarily describe the full causal dependence of fitness on genotype. There are (at least) three issues here. First, unless all the frequencies of all the alleles that affect fitness are included, the marginal selection on the alleles in the analysis will change as the causal genotypes change in frequency. Second, even if all causal alleles are included, the selection on individual alleles (that is, their marginal effect on fitness) will change as a result of dominance, epistasis and frequency-dependence. Third, unless recombination is much faster than other processes, the dynamics will involve associations amongst alleles (linkage disequilibria), and the marginal selection will depend on these associations as well as on the allele frequencies. If all genes that affect fitness are included, the multilocus selection coefficients will reflect the dependence of selection on these factors.

Some forms of frequency-dependent selection can be captured by considering the contribution of pairs, or larger clusters, of genes: this seems natural when the frequency dependence of genic fitness is due to the different frequency with which genes find themselves in diploid genotypes, formed by random union. This approach is traditional in game theory, which considers the payoff from interactions between two or more individuals. However, only polynomial frequency-dependence can be modelled in this way. Similar issues arise in coalescent theory, where selection can be represented, as we trace back through a genealogy, by the branching of two or more "virtual" ancestral lineages; once the ancestral selection graph is constructed, one of these branches is selected with a probability that depends on the genotype (Neuhauser, 1999). In this approach, k'th order frequency dependence is described by branching of $k + 1$ virtual ancestral lineages, just as k'th order polynomial frequency-dependence can be accommodated by considering groups of k interacting individuals. Thus, general (i.e., non-polynomial) frequency-dependence is intractable, both forwards and backwards in time.

Our arguments center on how selection coefficients should be defined and interpreted, given that the relation between genotype and fitness is known with certainty.

Thus, they are mainly applicable to theory rather than to data: measurement error, and ignorance of which loci actually cause fitness differences, would introduce extra layers of uncertainty. However, our theoretical arguments should help inform empirical research that aims at identifying the genes responsible for fitness differences, and distinguishing direct from indirect selection.

■ **Acknowledgements**

This work was supported by a grant from the European Research Council, ERC-2009-AdG-250152, to NB and a National Science Foundation grant, DEB 1255777, to MRS. We are most grateful for comments from Sumit Dhole, Denis Roze, Barbora Trubenova and Harold Vladar. We thank the referees for their helpful criticism.

References

Abbot, P., J. Abe, J. Alcock, S. Alizon, J. A. C. Alpedrinha, M. Andersson, J. B. Andre, M. van Baalen, F. Balloux, S. Balshine, et al. 2010. Inclusive fitness theory and eusociality. Nature 471:E1-E4.

Barton, N. H. 1983. Multilocus clines. Evolution 37:454-471.

Barton, N. H. 1986. The effects of linkage and density-dependent regulation on gene flow. Heredity 57:415-426.

Barton, N. H. 1995. A general model for the evolution of recombination. Genetical Research 65:123-144.

Barton, N. H., and M. Turelli. 1987. Adaptive landscapes, genetic distance, and the evolution of quantitative characters. Genetical Research 49:157-174.

Barton, N. H., and M. Turelli. 1991. Natural and sexual selection on many loci. Genetics 127:229-255.

Barton, N. H., and M. Turelli. 2006. Will population bottlenecks and multilocus epistasis increase additive genetic variance? Evolution 60:1763-1776.

Bürger, R. 1991. Moments, cumulants and polygenic dynamics. J. Math. Biol. 30:199-213.

Cheverud, J. M., and E. J. Routman. 1995. Epistasis and its contribution to genetic variance components. Genetics 139:1455-1461.

Christiansen, F. B. 1999. Population genetics of multiple loci. John Wiley & Sons, New York.

Dobzhansky, T. 1940. Speciation as a stage in evolutionary divergence. Amer. Nat. 74:312-321.

Feldman,M W,S P Otto,and F B Christiansen.1996. “Population Genetic Perspectives on the Evolution of Recombination.” Annual Review of Genetics 30:261–95.

Gardner, A., S. West, and N. H. Barton. 2007. The relation between multilocus population genetics and social evolution theory. American Naturalist 169:207-226.

Gardner, A., West, S.A. and Wild, G. 2011. The genetical theory of kin selection. J. Evol. Biol. 24: 1020-1043.

Hamilton, W. D. 1964a. The genetical evolution of social behaviour I. Journal of Theoretical Biology 7:1-16.

Hamilton, W. D. 1964b. The genetical evolution of social behaviour II. J. Theor.Biol. 7:17-52.

Kirkpatrick, M., T. Johnson, and N. H. Barton. 2002. General models of multilocus evolution. Genetics 161:1727-1750.

Kirkpatrick, M., M.J. Ryan. 1991. The evolution of mating preferences and the paradox of the lek. Nature 350:33-38.

Kirkpatrick, M., and M. R. Servedio. 1999. The reinforcement of mating preferences on an island. Genetics 151:865-884.

Neuhauser, C. 1999. The ancestral graph and gene genealogy under frequency-dependent selection. Theoretical Population Biology 56:203-214.

Nowak, M. A., C. E. Tarnita, and E. O. Wilson. 2010. The evolution of eusociality. Nature 466:1057–1062.

O'Donald, P. 1980. Genetic Models of Sexual Selection. Cambridge University Press, Cambridge, U.K.

Queller, D.C. 1992. A general model for kin selection. Evolution 46: 376-380.

Sober, E. 1984. The nature of selection: evolutionary theory in philosophical focus. MIT Press. Cambridge, Massachussetts.

Sober, E., and R. Lewontin. 1982. Artifact, cause and genic selection. Philosophy of Science 49:157-180.

Thomson, G. 1977. The effect of a selected locus on linked neutral loci. Genetics 85:753-788.

Turelli, M., and N. H. Barton. 1994. Genetic and statistical analyses of strong selection on polygenic traits: what, me normal? Genetics 138:913-941.

Figure 1

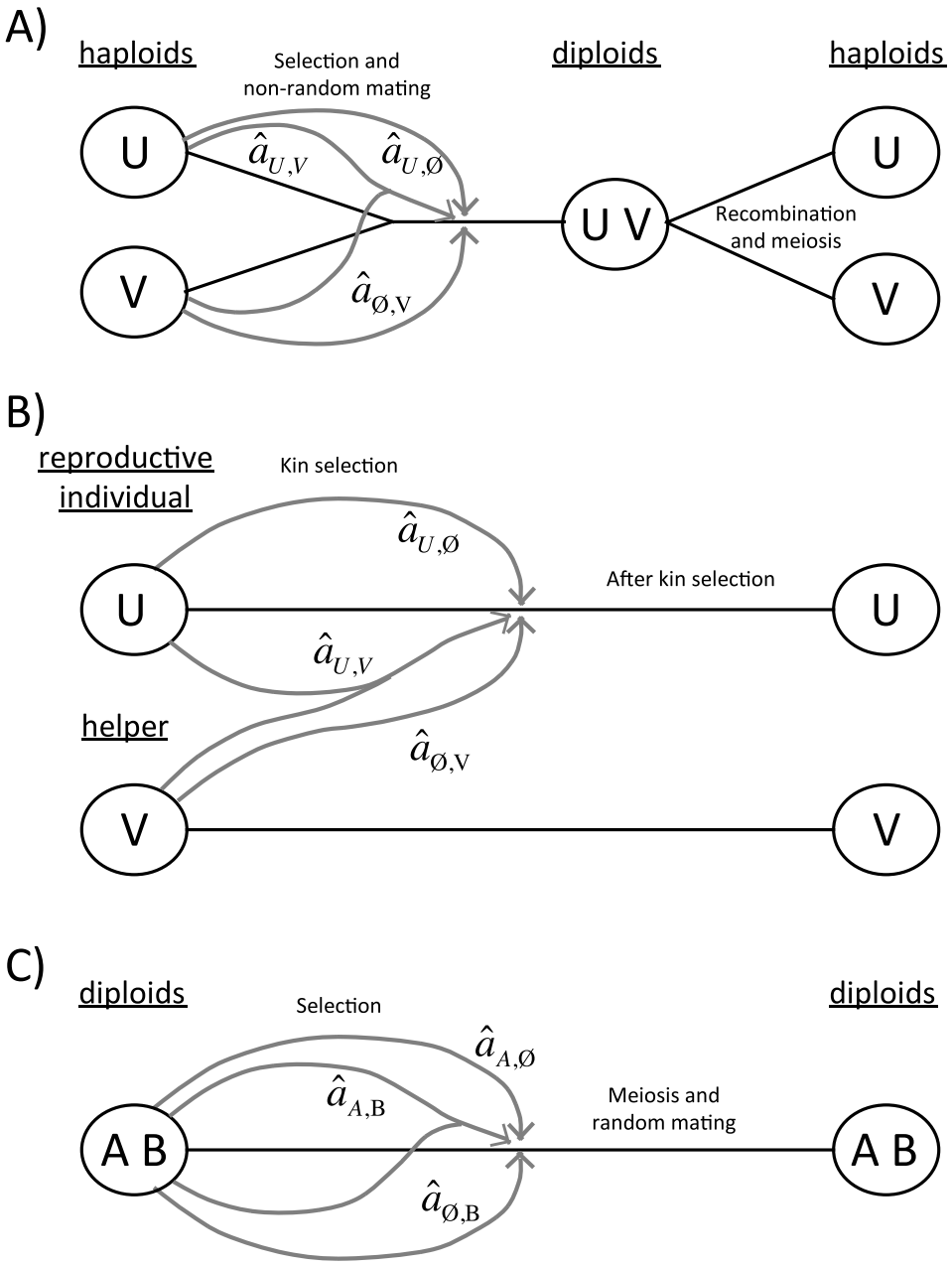

Figure 1. Life cycles and illustration of the action of the multilocus coefficients with selected examples, using the notation of Barton and Turelli (1991). Solid back lines link

points of interest in the life cycle, and dashed grey lines indicate the action of the selection coefficients. A) selection coeffcients $\hat{a}_{U,\emptyset}$ act on all subsets of the set U in females, selection coefficients $\hat{a}_{\emptyset,V}$ act on all subsets of the set V in males, and the coefficients $\hat{a}_{U,V}$ describe selection (including non-random mating) that brings together subsets of the sets U and V. B) Selection on the set of genes U in a particular individual comes from the genes it carries, $\hat{a}_{U,\emptyset}$, and from the set of genes V in helpers, $\hat{a}_{\emptyset,V}$, as well as from interactions between the two, $\hat{a}_{U,V}$. C) Selection coefficients can be written for the sets of loci A (representing additive effects), $\hat{a}_{A,\emptyset}$, and B (representing dominance effects), $\hat{a}_{\emptyset,B}$, as well as the interaction between them, $\hat{a}_{A,B}$.