Metabolic Models of Selection Response

PETER D. KEIGHTLEY†

Institute of Cell, Animal and Population Biology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, Scotland

Consequences of directional selection on metabolic flux are explored in models for which variation in flux among individuals is generated by segregation of allelic variants at enzyme activity loci. The pattern of selection response is strongly affected by the presence of genetic dominance and epistasis, which are automatically generated in metabolic systems. The expected magnitudes of dominance and epistasis effects on flux are evaluated. Small differences in enzyme activity generate little dominance, but a null allele will tend to be recessive for the pathway in which it occurs and for metabolically distant pathways. Epistasis is found to be greatest in short pathways in which large differences in enzyme activity occur. Under divergent artificial selection asymmetrical responses can occur due to the presence of directional dominance and epistasis, and lead to departures from the classic infinitesimal model of quantitative genetic variation. The effects of epistasis and dominance are in opposite directions, however, and partially cancel each other out in a diploid population.

© 1996 Academic Press Limited

Introduction

Quantitative genetic models rarely incorporate a biological model connecting variation of genes with variation in phenotypes. The most usual model of quantitative genetic variation is the infinitesimal model of many genes each with small additive effects (Fisher, 1918; Bulmer 1980). The infinitesimal model is widely used in practical selective breeding because it accurately predicts selection responses in the short term, and it is mathematically tractable. As an evolutionary model, it has severe limitations, however (Turelli & Barton, 1994). More realistic models of quantitative genetic variation would require knowledge of the nature of genes underlying quantitative traits. Metabolic fluxes and substrate concentrations are continuous characters under polygenic control. These “metabolic traits” are potentially influenced by all the enzymes of intermediary metabolism and represent a basic type of quantitative trait. While it can be argued that many of the traits studied by quantitative geneticists (e.g., components of fitness, behaviour, or bristle number in Drosophila) are highly complex with an indirect relationship to fluxes or metabolite pool levels, some traits that are the targets of selection by animal and plant breeders (e.g., growth rate, yield, fatness, or oil content) may be more closely related to metabolism.

Metabolic traits have several properties that are common phenomena for quantitative traits, for example, directional dominance (Kacser & Burns, 1981), epistasis (Dykhuizen et al., 1987; Keightley, 1989; Szathmary, 1993), and pleiotropy (Keightley & Kacser, 1987). It may be possible to improve the understanding of these phenomena by studying their basis in metabolism. Henrik Kacser believed it important to relate his work on consequences of changes of enzyme activities on systemic properties of metabolism to the response to artificial selection. This interest was stimulated by the environment of the Department of Genetics at Edinburgh, which is the main centre in the UK for quantitative genetics and research into animal breeding. During many informal discussions of the subject, Kacser proposed that if a trait is a linear function of flux, and genetic variation for enzyme activities generates genetic variation for

† E-mail: P.Keightley@edinburgh.ac.uk

© 1996 Academic Press Limited
flux, artificial selection will generate strongly asymmetrical responses to artificial selection because substitutions at different enzyme loci in metabolic systems tend to be synergistic for increases in enzyme activity and antagonistic for decreases in activity.

The purpose of this paper is to investigate selection responses in models of metabolic traits. Before considering the effects of selection on variation at several enzyme activity loci, it is necessary to consider the effects of single and double allelic substitutions on dominance and epistatic relationships, as these are important factors that can generate asymmetrical selection responses (Falconer, 1989).

Dominance

Kacser & Burns (1981) used Wright’s (1934) dominance index, \( D \), to describe the effect of an allelic substitution at an enzyme locus on the relationship between the three diploid phenotyped for a metabolic flux or metabolite concentration. The index is defined

\[
D = \frac{J_{WW} - J_{WM}}{J_{WW} - J_{MM}},
\]

where \( J_{WW} \), \( J_{WM} \) and \( J_{MM} \) are the measured phenotypes (e.g., fluxes) for the wild type, heterozygote and mutant respectively. A \( D \) value of zero occurs if the wild type and heterozygote phenotypes are identical, and implies complete recessivity of the mutant. A value of 0.5 occurs if the heterozygote phenotype is exactly intermediate between the mutant and wild type, and implies additive gene action.

Linear Pathway

In a linear pathway of enzymes catalysing monomolecular reactions the equation for the flux approximates to the following form if all the enzymes operate well below saturation:

\[
J \propto \frac{1}{(1/E_1 + 1/E_2 + \ldots + 1/E_n)}
\]

(Kacser & Burns, 1981), where \( E_1 \ldots E_n \) are terms proportional to the enzyme activities for the different steps. Under the assumption that the heterozygote enzyme activity is intermediate between the wild type and the mutant, Kacser & Burns showed that the dominance index is

\[
D = \frac{C_{Em}(1 - E_M/E_W) + E_M/E_W}{C_{Em}(1 - E_M/E_W) + E_M/E_W + 1},
\]

where \( E_W \) and \( E_M \) are the wild type and mutant enzyme activities respectively, and \( C_{Em} \) is the control coefficient of flux with respect to the wild type enzyme activity. For the case of a null or nearly null mutant enzyme activity (\( E_M/E_W \to 0 \)), and a small control coefficient of the wild type enzyme activity (\( C_{Em} \to 0 \)), \( D \) approaches zero, implying the mutant phenotype is recessive. For “quantitative variants” with small differences in activity between “wild type” and “mutant” (\( E_M/E_W \approx 1 \)), the expected value for \( D \) is 0.5 irrespective of the control coefficient, and implies additive gene action.

Branched Pathways

For branched pathways in which competition for a common substrate at a branch point occurs (Fig. 1), variation of an enzyme activity in any of the three branches affects fluxes in the common and both of the competing branches. If the measured phenotype is the flux (e.g. \( J_B \)) in the branch in which enzyme activity variation occurs, it can be shown that the equation for flux is

\[
J \propto \frac{1}{1/E + k + x},
\]

where \( E \) is proportional to the enzyme activity in question, \( k \) is a sum of terms inversely proportional to the enzyme activities in the branch other than \( E \), and

\[
x = \frac{AC}{A + C/K_{AS}}.
\]

\( A \) and \( C \) are “group enzyme activities”, i.e., sums of terms inversely proportional to enzyme activities in the common \( (A) \) and competing \( (C) \) branches respectively, and \( K_{AS} \) is the overall equilibrium constant for the reactions in the common branch. The term \( x \) is therefore always positive. Equation (4) is of

\[
\text{Fig. 1. A branched metabolic pathway with a common (A) and two competing branches (B and C). Each branch is a chain of monomolecular reactions catalysed by enzymes far below saturation. Competition for the common substrate, S, occurs at the branch point.}
\]
a similar form to the equation for the linear chain (2), so it follows that the dominance index is also of a similar form to that for the straight chain (3). The control coefficient of flux in a branch with respect to an enzyme in the same branch is

$$C = \frac{1/E}{1/E + k + \alpha},$$  

(6)

and it therefore follows that the sum of control coefficients of the flux in a branch with respect to the enzymes in that branch is less than unity.

An interesting characteristic of systems with monomolecular reactions far below saturation is that the dominance indices for all fluxes or metabolite concentrations in the system are the same, irrespective of the complexity of the system (Keightley & Kacser, 1987). For example, in the branched pathway, enzyme variation in the common branch will generate the same dominance indices for either of the output fluxes. If there are several enzymes in a branch, or if the $\alpha$ term in (6) is large relative to $1/E$, the arguments used for the isolated straight chain also apply to the branched pathway. Recessivity for null mutants and additivity for “quantitative” variants are expected for the flux in the branch and all the other fluxes or substrate concentrations in the system. This is surprising in view of the fact that control coefficients of fluxes with respect to enzymes in competing pathways can have absolute values greater than unity. Equality of dominance indices for different traits implies that a flux phenotype could show recessivity for a mutant occurring in a metabolically distant pathway, in which case the additive effect of the mutant would be small. Differences in dominance for different metabolic traits can occur with nonlinear enzymes (Keightley & Kacser, 1987), and can also lead to dominance for null alleles (Cornish-Bowden, 1987).

**Epistasis**

Genetic epistasis is a change in the effect of an allelic substitution on the phenotype due to an allelic substitution at a different locus. In diploids, epistasis is complicated by the fact that interaction can occur between the average effects of substitutions (generating additive by additive epistasis), dominance effects (generating dominance by dominance epistasis), or between average effects and dominance effects (generating dominance by additive epistasis) (Falconer, 1989, Chapter 7). In a haploid, the complication of dominance is not present. Consider a haploid model with two variable enzymes each with two alleles in a linear pathway of enzymes far below saturation catalysing monomolecular reactions. An “interaction index” measuring the extent of epistasis for flux can be defined as

$$I = \frac{J_{HH} - J_{LL}}{(J_{HL} - J_{LL}) + (J_{HH} - J_{LL})},$$  

(7)

where $J_{HH}$ is the flux for the genotype with alleles producing high enzyme activities at both variable loci, $J_{LL}$ is the flux for the low enzyme activity genotype, and $J_{HL}$ and $J_{HH}$ are fluxes for genotypes with one high and one low activity enzyme. A similar interaction index was described by Burns (1971), and an analysis of the consequences of changes of several enzyme activities at a time in terms of flux “deviation indices” carried out by Small & Kacser (1993). The interaction index is the ratio of the difference in flux generated by simultaneous double substitution to the sum of the differences generated by individually substituting one low activity allele for a high activity allele. If effects are additive, $I = 1$. Values less than unity imply antagonistic epistasis while values greater than unity imply synergistic epistasis. It turns out that epistasis is always synergistic as long as simultaneous substitutions at enzyme activity loci from low to high activity alleles are considered. Conversely, epistasis is antagonistic for high to low substitutions. [If the measured phenotype is a nonlinear function of flux, epistasis can be synergistic for high to low substitutions (Szathmary, 1993).] Under the simplifying assumption that the two variable enzymes have equivalent activities such that $J_{HL} = J_{HH}$ (to make the algebra tractable), and substituting the equation for the flux in the linear pathway in terms of the high and low enzyme activities ($E_H$ and $E_L$ respectively), the equation for the index becomes

$$I = \frac{1/E_L + 1/E_H + k}{2/E_H + k} = \frac{C_{HH}}{C_{HL}},$$  

(8)

where $k$ is the sum of reciprocals of terms proportional to the enzyme activities other than $E$, $C_{HH}$ is the control coefficient of the flux with respect to the high enzyme activity in the high allele background, and $C_{HL}$ is its control coefficient in the background of the low allele. Equation (8) illustrates the way in which the amount of interaction depends on how the control coefficients of flux change as a consequence of change in the genetic background from a low to a high activity allele. Interaction is at a maximum ($I \to \infty$) for null alleles, a low allele generates a blocked pathway and $C_{HL}$ is zero. Values of $I$ close to unity imply a lack of epistasis, and can occur under two circumstances: (1) if the control of flux is dominated by enzymes in the pathway other
than the two variable loci on both high and low backgrounds, i.e., $k \gg 1/E_L$ and $k \gg 1/E_H$; (2) if there is only a small difference between the high and low enzyme activities, i.e., $E_L/E_H \rightarrow 1$. Values of $I$ as a function of $E_L/E_H$ are shown in Fig. 2 for pathways with different numbers of steps catalysed by enzymes with equivalent activities. Interaction is greater in shorter pathways because the $k$ term above becomes smaller. For pathways with more than, say, five steps, little interaction is present for $E_L/E_H$ values greater than 0.5, so quite large differences in enzyme activity between high and low alleles will often generate little interaction. However, interaction will increase as the number of variable enzymes increases. The extent of interaction and dominance variance for a population for a quantitative trait can also be quantified for models in which metabolic traits are affected by segregation of enzyme activity alleles (Keightley, 1989). The amount of interaction variance is highest with high frequencies of mutant alleles, in contrast to the dominance variance, which is highest when mutant alleles are rare, and this affects the pattern of responses of metabolic traits to directional selection (see below).

A practical investigation of epistasis has recently been carried out by Niederberger et al. (1992). Large simultaneous increases in several enzyme activities controlling flux in the tryptophan biosynthesis pathway of yeast were shown to generate a large amount of synergistic epistasis. Epistasis for changes in enzyme activities of the lactose operon of *E. coli* was also found by Dykhuizen et al. (1987).

### Response to Selection

Asymmetrical responses to artificial selection are often seen in breeding experiments, but there are several possible explanations (Falconer, 1989). Let us explore the consequences of directional selection on flux in haploid and diploid models. Assume that the phenotype under selection is the flux in a linear pathway, and alleles generating variability for enzyme activities segregate in an infinite population. For upward selection on flux, assume that the fitness of an individual, $W$, is a function of its flux according to

$$W = 1 + (J - J_0)/J_0,$$

where $J_0$ is the mean flux phenotype in the base population before commencement of selection. For downward selection the fitness is

$$W = 1 - (J - J_0)/J_0.$$
accelerating pattern of response for upward selection. The response attenuates as allele frequencies approach fixation. For downward selection epistasis is antagonistic and the response attenuates much quicker without a period of acceleration. In cases of small differences in enzyme activity response is closer to symmetrical because the genetic variation is mostly additive (cf. Fig. 2).

DIRECTIONAL SELECTION IN A DIPLOID MODEL

Examples of the pattern of response to selection with a model the same as above except diploid are shown in Fig. 4. The shape of the response curves are qualitatively similar to the haploid case, but less asymmetrical and with a less sigmoidal shape for upward selection. In a diploid population directional genetic dominance as well as epistasis are present, and these act in opposite directions to one another. The effect of epistasis is similar to the haploid model, i.e., synergistic upwards and antagonistic downwards. Upward selection implies selection of genetically dominant alleles and leads to a smaller response that selection downwards on recessive alleles at intermediate frequencies (Falconer, 1989). The extent to which the selection response departs from the infinitesimal model of many additive genes of small effect (the model most commonly assumed in quantitative genetics) expressed as the percentage by which the observed response is greater or less than that expected under the infinitesimal model is shown in Table 1.

Table 1

<table>
<thead>
<tr>
<th>$E_t/E_h$</th>
<th>Generation</th>
<th>Up selection</th>
<th>Down selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>1</td>
<td>6</td>
<td>−5</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>29</td>
<td>−57</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>3</td>
<td>−73</td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>2</td>
<td>−8</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>1</td>
<td>−17</td>
</tr>
</tbody>
</table>

Variation for flux occurs in an infinite population and is genetically determined by segregation of two alleles at an initial frequency of 0.5 at each of ten enzyme activity loci acting in a metabolic pathway. The figures are the percentage difference in response observed with the metabolic model at different generations from the response expected under the infinitesimal model with equivalent additive genetic variance at generation 0.

Stabilising selection in natural populations

Finally, let us consider the consequences for selection for an intermediate optimum. Many traits are thought to be subject to this type of selection in nature (Endler, 1986). The consequences of stabilising selection under additive models have been worked out (Turelli, 1984; Keightley & Hill, 1988; Burger & Lande, 1994), but recently effects of stabilising selection on metabolic traits have received attention. Beaumont (1988) investigated the effect on variability of metabolic traits generated by mutations that reduce enzyme activities. While reduced flux is the likely outcome, effects on substrate levels are less easy to predict and depend on whether the mutant enzyme is “upstream” or “downstream” in the pathway from the substrate. If mutations randomly occur at loci controlling different enzyme activities in a pathway, intermediate optima for substrate levels may be common. Clark (1991) also hypothesised that substrate pool levels may often be subject to stabilising selection. Under the assumption of an optimal model for flux or substrate level, there will be a plateau in the fitness function, and Clark showed that the expected equilibrium distribution of allelic
effects in a population will be skewed upwards. This equilibrium distribution will also depend on the distribution of effects of new mutations on enzyme activities, for which the only available information suggests distributions that depart strongly from normality (Clark, et al., 1995). Further data on the distribution and frequencies of effects of enzyme activity alleles segregating in natural populations and distributions of effects of new mutations on enzyme activities has the potential to shed further light on the nature of quantitative genetic variation for metabolic traits.

I dedicate this paper to Henrik Kacser. Much of the work described here is based on a Genetics Honours project that I carried out under Henrik’s guidance in 1981–82. I fondly remember having many hours of stimulating debate with him on this topic. I thank W. G. Hill, J. R. Small and A. Cornish-Bowden for helpful comments on the manuscript.

REFERENCES
