Available online at www.sciencedirect.com
SCIENCE DIRECT®
Theoretical Population Biology 63 (2003) 105–114
Theoretical Population Biology
http://www.elsevier.com/locate/ytpbi

The balance between pleiotropic mutation and selection, when alleles have discrete effects

D. Waxman*

Centre for the Study of Evolution, School of Biological Sciences, University of Sussex, Brighton BN1 9QG, Sussex, UK

Received 26 March 2002

Abstract

The theory of pleiotropic mutation and selection is investigated and developed for a large population of asexual organisms. Members of the population are subject to stabilising selection on Ω phenotypic characters, which each independently affect fitness. Pleiotropy is incorporated into the model by allowing each mutation to simultaneously affect all characters. To expose differences with continuous-allele models, the characters are taken to originate from discrete-effect alleles and thus have discrete genotypic effects. Each character can take the values $n \times \Delta$ where $n = 0, \pm 1, \pm 2, \ldots$, and the splitting in character effects, Δ, is a parameter of the model. When the distribution of mutant effects is normally distributed around the parental value, and Δ is large, a "stepwise" model of mutation arises, where only adjacent trait effects are accessible from a single mutation. The present work is primarily concerned with the opposite limit, where Δ is small and many different trait effects are accessible from a single mutation.

In contrast to what has been established for continuous-effect models, discrete-effect models do not yield a singular equilibrium distribution of genotypic effects for any value of Ω. Instead, for different values of Ω, the equilibrium frequencies of trait values have very different dependencies on Δ. For $\Omega = 1$ and 2, decreasing Δ broadens the width of the frequency distribution and hence increases the equilibrium level of polymorphism. For all sufficiently large values of Ω, however, decreasing Δ decreases the width of the frequency distribution and the equilibrium level of polymorphism. The connection with continuous trait models follows when the limit $\Delta \to 0$ is considered, and a singular probability *density* of trait values is obtained for all sufficiently large Ω.

Keywords: Pleiotropy; Mutation selection balance; Stabilising selection; Quantitative traits; Asexual population

1. Introduction

It is a common strategy of evolutionary theorists to make the following assumptions about a population under consideration. (i) The population is fully characterised by a series of phenotypic traits. (ii) Each trait has an optimum value. (iii) The optimum value of each trait is independent of the state of the other traits.

The above, although clearly a gross over-simplification of the evolutionary process, is however thought to capture some of the essentials of the process of natural selection. It is the basis of otherwise different modelling approaches such as Fisher's geometrical model (Fisher, 1930; Barton, 1998; Orr, 1998, 2000; Barton and Keightley, 2002) or quantitative genetic models of multiple traits under some form of selection (e.g. Lande, 1980; Turelli, 1985).

Although selection may act independently on the traits, a common assumption is that all traits are controlled, to some extent, by the same genetic sequence, and as such, gene action is pleiotropic. This reflects the ubiquity of pleiotropic gene action in nature (Caspari, 1952; Wright, 1968).

Although empirical evidence and theoretical arguments suggest that much of the stabilising selection measured in the laboratory or field may be merely "apparent" (the result of deleterious pleiotropic effects being associated with extreme trait values (Robertson, 1967; Barton, 1990; Keightley and Hill, 1990; Nuzhdin et al., 1999))—the simple model described above remains compelling for qualitative insight if not quantitative prediction of particular cases.

A series of important questions about these models must concern the availability of different alleles. The

*Fax: +44-1273-678433.
E-mail address: d.waxman@sussex.ac.uk.

0040-5809/03/$ - see front matter
doi:10.1016/S0040-5809(02)00036-9

most popular modelling strategy is to extend Crow and Kimura's (1964) continuum-of-alleles model to multiple dimensions, and assume that the mutational changes in different characters are statistically independent (Lande, 1980; Turelli, 1985; Waxman and Peck, 1998, 2000). The number of possible alleles that have to be posited in order to make these multidimensional analyses plausible is, however, extremely large in sexual populations. Using a small number of alleles, however, as in e.g. Turelli's (1985) five-allele model, automatically places constraints on the pleiotropy. Wagner (1989), investigated a model that utilised a continuum-of-alleles, but with a rectangular matrix relating the vector of "trait values" to the vector of "allelic effects", thereby constraining all possible allelic effects at a given locus to a line in "trait space". As pointed out by Wagner, such constraints affect the results of the analyses irrespective of the number of alleles. A problem with this sort of model is that, although we must assume that some such constraints exist, we have little principled way of determining what they might be. The empirical evidence available suggests that Wagner's linear approximation may be over-simple. In model systems, different alleles have been detected with very different sets of effects on multiple traits (Barton and Turelli, 1989; Lyman and Mackay, 1998).

In the present paper, we address exclusively the case of a population of asexual organisms. In this case, the entire genome can be thought of as a single haploid locus, with a very large number of different possible alleles. This allows the notion of unconstrained pleiotropy, where the effects of an allele on different traits are statistically independent.

A key feature of models with a continuum of alleles, is that when the number of traits becomes sufficiently large, the equilibrium distribution of genotypic effects is singular. The nature of the singularity is the presence of an infinitely narrow spike of finite weight—a Dirac delta function (Waxman and Peck, 1998, 2000). This spike corresponds to a non-negligible proportion of the population having genotypic trait values lying *exactly* at the optimal value—despite the existence of continuous-effect mutations and continuous-effect traits. The singular spike is present whenever Ω is equal or larger than a critical value, termed Ω_c and the smallest possible value of Ω_c is $\Omega_c = 3$ (Waxman and Peck, 1998, 2000). Singular equilibrium behaviour would occur in the absence of mutations, for any Ω, since the population would have been driven, by selection, to the fitness optimum, with the population possessing no variation around this value. The singular behaviour persists even in the presence of mutations, for $\Omega \geqslant \Omega_c$, since a sharp reduction occurs in the proportion of mutations that can take non-optimal genotypes to optimal or near-optimal values (Waxman and Peck, 2000).

At first sight, any work on this subject does appear to have a direct bearing on the maintenance of genetic variation. However, the presence or absence of a sharply peaked distribution of genotypic effects in the vicinity of the mean trait values has very little effect on the genetic variance. As noted in Waxman and Peck (1998), however, features of the shape of the distribution of genotypic effects may have implications for other significant properties such as conservation of protein sequences in widely diverged populations, despite the fact that mutations of small effect may occur and selection need not be strong.

This work aims to address what happens when allelic effects, and hence trait effects, are not continuous, so the frequency of any particular (discretely labelled) genotype cannot be greater than unity. In such a case, there cannot be even the notion of a singular distribution, yet in the related continuous trait problem, there is a singular distribution for all $\Omega \geqslant \Omega_c$.

Our model differs from the discrete allele model investigated by Slatkin (1987) and Slatkin and Frank (1990), in that those models included "stepwise" mutation schemes in which only adjacent mutation types are accessible from a single mutation. The present investigation sheds light into what occurs when alleles have discrete effects, but are not restricted to short-range stepwise-type models.

2. Non-pleiotropic haploid dynamics

Consider an effectively infinite population of asexual individuals for whom stabilising selection occurs on the value of a single quantitative trait. Let us assume mutations occur with sufficiently low probability, that there is negligible chance of an offspring containing two or more new mutations. Then individuals may be mathematically treated as a single haploid locus. Let possible genotypes of individuals be discrete and denoted by T_n with $n = 0, \pm 1, \pm 2, \ldots$. The genotypic effect of a T_n genotype individual is $n \times \Delta$, where Δ is a parameter characterising the model that we shall refer to as the "splitting in genotypic or trait values." Censusing in the juvenile phase, immediately after birth, the events in the discrete generation lifecycle are viability selection, followed by the birth of all offspring, which in turn is followed by the death of all parents. Fertility is assumed to be independent of genotype and mutations are taken to occur at birth.

Let p_n denote the frequency of individuals with genotype T_n in one particular generation. In the next generation, we have

$$p_n' = \frac{\sum_m M_{nm} w_m p_m}{\sum_m w_m p_m}, \tag{1}$$

where here, and throughout this work, we adopt the convention that summations with unspecified limits have a summation index that runs over its entire infinite range. In (1), w_m is proportional to the viability of individuals with genotype T_m. Assuming the phenotypic value of an individual is the sum of a genotypic value and an independent random environmental contribution, the quantity w_m follows from an average of viability, as a function of phenotypic value, over environmental effects (see e.g. Turelli, 1984). Following convention, w_m is defined so that its maximum value is unity. The quantity M_{nm} is the probability that an offspring produced by a T_m genotype individual will be of T_n genotype. The probability that any offspring has one of the possible genotypes is unity, hence $\sum_n M_{nm} = 1$.

Let u be the probability of a mutation occurring per generation. Then the probability of an offspring containing no mutations is $M_{nn} = 1 - u$. For $n \neq m$, we write

$$M_{nm} = u \times f_{nm}, \quad n \neq m \tag{2}$$

and $\sum_{n(\neq m)} f_{nm} = 1$ because $\sum_n M_{nm} = 1$. We then have $p'_n = (\sum_m w_m p_m)^{-1}[(1-u)w_n p_n + u\sum_{m(\neq n)} f_{nm} w_m p_m]$. Under weak selection, w_n differs little from unity and we take a quadratic fitness model (Kimura, 1965): $w_n = 1 - sn^2\Delta^2$, where s is a positive constant whose size is a measure of the strength of selection. Omitting small terms that are higher than first order in u, the selection coefficient $sn^2\Delta^2$ and any products of these, leads to

$$p'_n \simeq (1 - u - sn^2\Delta^2 + s\mu_2)p_n + u\sum_{m(\neq n)} f_{nm} p_m, \tag{3}$$

where $\mu_2 = \sum_n n^2\Delta^2 p_n$ is the mean square value of the trait.

We also follow Kimura (1965) in taking a Gaussian distribution of mutant effects; thus,

$$f_{nm} = f(n-m) = \frac{e^{-(n-m)^2\Delta^2/(2\sigma^2)}}{\sum_{m(\neq 0)} e^{-m^2\Delta^2/(2\sigma^2)}}, \quad n \neq m, \tag{4}$$

where σ^2 is measure of the width of the mutation distribution. Note that when $\Delta/\sigma \ll 1$, the variance of f_{nm} is $\sum_{n(\neq m)}(n-m)^2 f(n-m) \simeq \sigma^2/\Delta^2$ while if $\Delta/\sigma \gg 1$, the variance of f_{nm} is approximately 1. Note also that when $\Delta/\sigma \gg 1$, $f_{nm} \simeq 1/2(\delta_{n,m+1} + \delta_{n,m-1})$, where $\delta_{n,m} = 1$ if $n = m$ and is zero otherwise. This limiting form of f_{nm} corresponds to the stepwise model of mutation used by Slatkin (1987), where only neighbouring trait effects are reachable from a single mutation.

Restricting all further considerations to a description of the various frequencies p_n at equilibrium, it follows that $p'_n = p_n$ and (3) can be rewritten in the form

$$n^2 p_n - \frac{u}{s\Delta^2}\sum_{m(\neq n)} f_{nm} p_m = -\alpha p_n, \tag{5}$$

where

$$\alpha = (u - s\mu_2)/(s\Delta^2). \tag{6}$$

Eq. (5) is of the form of an eigenvalue equation where $-\alpha$ is the eigenvalue and $L_{nm} = n^2\delta_{n,m} - u/(s\Delta^2)f_{nm}(1 - \delta_{n,m})$ are the elements of a linear operator L. A consideration of the dynamical equation, (3), indicates that $-\alpha$ is, in fact, the smallest eigenvalue of L and the eigenvector of L belonging to this eigenvalue corresponds to the set of equilibrium frequencies, p_n. The values of α and the p_n can be determined numerically.

3. Generalisation to models with pleiotropy

Let us now consider pleiotropic models where fitness is determined by a number of phenotypic traits. Assume there are Ω such traits and these have genotypic effects labelled by $n_1, n_2, \ldots, n_\Omega$ (where $n_i = 0, \pm 1, \pm 2, \ldots$ and $i = 1, 2, \ldots, \Omega$) and the corresponding genotypic values on the Ω traits are $n_1\Delta, n_2\Delta, \ldots, n_i\Delta$.

As stated in the Introduction, all Ω traits are assumed to be controlled by the same genetic sequence, so that when a mutation occurs, all traits generally change. We further assume, as stated in the Introduction, that the mutational effects on the Ω traits are statistically independent, since this is a plausible hypothesis that is consistent with at least some of the experimental data (Lyman and Mackay, 1998). Thus, the appropriate generalisation of the previous section involves replacing single indices such as n (which label trait values), by column vectors with Ω components that label the trait values on all Ω characters. Vectors are denoted by a bold symbol e.g. $\mathbf{n} = (n_1, n_2, \ldots, n_i)^{\mathrm{T}}$, where T denotes transposition, and the Euclidean length of $\mathbf{n}$ is the positive quantity $||\mathbf{n}||$, defined as $||\mathbf{n}|| = \sqrt{n_1^2 + n_2^2 + \cdots + n_i^2}$.

With $\mathbf{0} = (0, 0, \ldots, 0)^{\mathrm{T}}$, the distribution of mutant effects is given by

$$f_{\mathbf{nm}} = f(\mathbf{n} - \mathbf{m}) = \frac{\prod_{j=1}^{\Omega} e^{-(n_j - m_j)^2\Delta^2/(2\sigma^2)}}{\sum_{\mathbf{m}(\neq \mathbf{0})}\prod_{j=1}^{\Omega} e^{-m_j^2\Delta^2/(2\sigma^2)}} = \frac{e^{-||\mathbf{n}-\mathbf{m}||^2\Delta^2/(2\sigma^2)}}{\sum_{\mathbf{m}(\neq \mathbf{0})} e^{-||\mathbf{m}||^2\Delta^2/(2\sigma^2)}}. \tag{7}$$

Let us assume all Ω traits affect fitness independently, so $w_{\mathbf{n}} = \prod_{i=1}^{\Omega}(1 - sn_i^2\Delta^2)$. The quantity s is a measure of the strength on selection on any single trait and is directly analogous to the quantity s in Eq. (5). Under weak stabilising selection, this approximates to a fitness

function with additive selection coefficients

$$w_{\mathbf{n}} \simeq \left(1 - \sum_{i=1}^{\Omega} s n_i^2 \Delta^2\right) = 1 - s||\mathbf{n}||^2\Delta^2 \tag{8}$$

and Eqs. (4) and (5) yield

$$||\mathbf{n}||^2 p_{\mathbf{n}} - \frac{u}{s\Delta^2} \sum_{\mathbf{m}(\neq\mathbf{n})} f_{\mathbf{nm}} p_{\mathbf{m}} = -\alpha p_{\mathbf{n}} \tag{9}$$

with α given by (6) but now $\mu_2 = \sum_{\mathbf{n}} \Delta^2 ||\mathbf{n}||^2 p_{\mathbf{n}}$.

4. Analysis for $\Delta/\sigma \ll 1$

We first briefly note the results when $\Delta/\sigma \gg 1$, before dealing at greater length with the case $\Delta/\sigma \ll 1$.

When $\Delta/\sigma \gg 1$, we have that $f_{\mathbf{nm}} \simeq (2\Omega)^{-1}$ if $||\mathbf{n}-\mathbf{m}|| = 1$ and $f_{\mathbf{nm}}$ is approximately zero for other values of $||\mathbf{n}-\mathbf{m}||$. This is the direct extension of the stepwise mutation model to higher than one dimension. For example if $\Omega = 2$, $f_{\mathbf{nm}} \simeq (\delta_{n_1+1,m_1}\delta_{n_2,m_2} + \delta_{n_1-1,m_1}\delta_{n_2,m_2} + \delta_{n_1,m_1}\delta_{n_2+1,m_2} + \delta_{n_1,m_1}\delta_{n_2-1,m_2})/4$. This model may be analysed very simply in terms of a non-pleiotropic stepwise mutation model: see Appendix A. Thus, in this case pleiotropy does not add a new aspect to the problem. More complicated stepwise schemes of mutation that do not, apparently, follow from the limit of a single Gaussian distribution of mutant effects have also been investigated by Slatkin and Frank (1990). The analysis presented here contrasts with these previous analyses by considering the regime $\Delta/\sigma \ll 1$. This corresponds to a very small splitting of trait values and leads to a distribution of mutant effects that connects trait categories with widely different labels, $\mathbf{n}$.

To proceed, note, using the Poisson summation method (Apostal, 1979), that we can write the denominator of (7) in the form $\sum_{\mathbf{m}(\neq\mathbf{0})} e^{-||\mathbf{m}||^2\Delta^2/(2\sigma^2)} = \left(\sqrt{\frac{2\pi\sigma^2}{\Delta^2}}\sum_k \exp\left[-\frac{(2\pi k)^2\sigma^2}{2\Delta^2}\right]\right)^{\Omega} - 1$. Thus since $\Delta/\sigma \ll 1$, the only term in the sum over k that is appreciable is the one with $k = 0$, and $\sum_{\mathbf{m}(\neq\mathbf{0})} e^{-||\mathbf{m}||^2\Delta^2/(2\sigma^2)} \simeq (2\pi\sigma^2/\Delta^2)^{\Omega/2} - 1 \simeq (2\pi\sigma^2/\Delta^2)^{\Omega/2}$. We can thus write (9) for $\Delta \ll \sigma$ as

$$||\mathbf{n}||^2 p_{\mathbf{n}} - \kappa_\Omega \sum_{\mathbf{m}(\neq\mathbf{n})} e^{-||\mathbf{n}-\mathbf{m}||^2\Delta^2/(2\sigma^2)} p_{\mathbf{m}} = -\alpha p_{\mathbf{n}}, \tag{10}$$

where

$$\kappa_\Omega \simeq \frac{u}{s}\frac{\Delta^{\Omega-2}}{(2\pi\sigma^2)^{\Omega/2}}. \tag{11}$$

The parameter s appearing in (11) is, by (8), the strength of selection on any single trait. To make a meaningful comparison of different splittings of trait effects, Δ—for different degrees of pleiotropy—we need to choose s so the combined effect of selection on all traits has, by an appropriate measure, an overall strength that is virtually independent of the degree of pleiotropy. From Section 2, we note that the standard deviation of mutant effects on any single trait is, for $\Delta \ll \sigma$, given by σ^2. Since mutation acts independently on all Ω traits, it follows that the variance of mutant effects, summed over all traits, is $\Omega\sigma^2$. Thus, unless the strength of selection is appropriately scaled, single mutant offspring of optimal or near-optimal individuals in models with pleiotropy ($\Omega > 1$) will, on average, suffer larger selection coefficients against them than corresponding offspring in non-pleiotropic models ($\Omega = 1$). A natural way to measure the overall strength of selection is as the mean fitness of single-mutant offspring of optimal fitness individuals. This quantity is $\sum_{\mathbf{n}} w_{\mathbf{n}} f_{\mathbf{n0}}$ and we shall adjust it to be effectively independent of the degree of pleiotropy, Ω. We find $\sum_{\mathbf{n}} w_{\mathbf{n}} f_{\mathbf{n0}} = 1 - s\sigma^2\Omega[1 + O((\Delta/\sigma)^{\Omega})]$. Since our considerations are restricted to the regime $\Delta/\sigma \ll 1$, to lowest order in Δ/σ, we have $\sum_{\mathbf{n}} w_{\mathbf{n}} f_{\mathbf{n0}} = 1 - s\sigma^2\Omega$. It then follows that if

$$s = s_0/\Omega \tag{12}$$

with s_0 a constant that is independent of Ω and Δ, then $\sum_{\mathbf{n}} w_{\mathbf{n}} f_{\mathbf{n0}}$ is effectively independent of the degree of pleiotropy. In what follows, we take s to be given by (12).

Returning to (10), let us note that amongst other things, it allows us to establish that all $p_{\mathbf{n}}$ are non-zero. This follows from the non-vanishing of $e^{-||\mathbf{n}-\mathbf{m}||^2\Delta^2/(2\sigma^2)}$ and positiveness of α. This last fact follows from contradictions that arise if we assume α is zero or negative. For example, if $\alpha = 0$, setting $\mathbf{n} = \mathbf{0}$, in (10), indicates that the only possible solution is $p_{\mathbf{0}} = 1$ and all other frequencies vanish. This leads to a contradiction, when we choose $\mathbf{n} \neq \mathbf{0}$ in the same equation. Similarly, choosing α negative leads to contradictions when again we set $\mathbf{n} = \mathbf{0}$.

Let us now make some general observations about (10). The term $||\mathbf{n}||^2 p_{\mathbf{n}}$ on the left-hand side arises from selection, and gives a quantitative measure of how selection influences the various discrete categories of trait effects. The other term on the left-hand side, involving the factor κ_Ω, arises from mutation and we can view κ_Ω as an effective mutation rate of the various discrete categories of trait effects. The balance between the opposing evolutionary forces of selection and mutation determine the equilibrium pattern of the $p_{\mathbf{n}}$'s. If selection is relatively strong compared with mutation, then we expect only a small number of $p_{\mathbf{n}}$'s will have appreciable values, and these will lie in the vicinity of $\mathbf{n} = 0$. By contrast if mutation is relatively strong compared with selection, then a much broader pattern of $p_{\mathbf{n}}$'s is expected.

Unlike the continuous trait calculations of Waxman and Peck (1998, 2000), no transition occurs, in a discrete

effect model, from a non-singular to a singular distribution at the critical value of $\Omega = \Omega_c$ (which for small mutation rates, is $\Omega_c = 3$). This absence of a transition applies, since all of the $p_{\mathbf{n}}$ are frequencies (i.e. probabilities), and hence can only lie in the range $\mathbf{1} \geqslant p_{\mathbf{n}} \geqslant 0$. Indeed because of this, there cannot be even the notion of a transition to a singular distribution. There is, however a change-over in behaviour that distinguishes $\Omega \geqslant \Omega_c$ from smaller values of Ω and this is the analogue of the transition of continuous models, that exists when trait effects are discrete.

The change-over in behaviour manifests itself in the way the $p_{\mathbf{n}}$ respond to changes in Δ. We shall consider what happens when we decrease Δ, and this involves following the detailed implications of this decrease.

When $\Omega = 1$, decreasing Δ has the effect of *increasing* κ_1, since $\kappa_1 \propto \Delta^{-1}$. If this were the only dependence on Δ, the increase in the effective mutation rate, κ_1, would automatically result in a broadening of the distribution of the $p_{\mathbf{n}}$'s, when plotted against n, the label of the discrete trait categories. There is, however, additional Δ dependence in (10). This is in the term originating from mutation, and containing what is effectively, the distribution of mutant effects of the different trait categories, $e^{-||\mathbf{n}-\mathbf{m}||^2\Delta^2/(2\sigma^2)}$. The effect on this term, of decreasing Δ, is to *increase* the range of trait categories accessible via a mutation. Thus, decreasing Δ in $e^{-||\mathbf{n}-\mathbf{m}||^2\Delta^2/(2\sigma^2)}$ also has the tendency to broaden the distribution of the $p'_{\mathbf{n}}$**s**. Thus for $\Omega = 1$, decreasing Δ has the overall effect of broadening the distribution of the $p_{\mathbf{n}}$'s and corresponds to increasing the equilibrium level of polymorphism. An example of this is shown in Fig. 1.

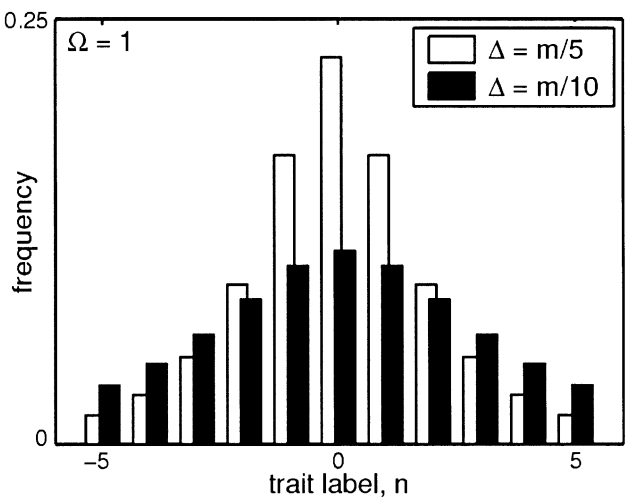

Fig. 1. For $\Omega = 1$ (the non-pleiotropic model), the equilibrium frequencies of different discrete categories of trait effects, p_n, are plotted against the trait label, n. The p_n were determined from numerical solution of (10), as outlined in Appendix B. The parameter values adopted are $s_0 = \frac{1}{40}$, $m = 0.2$, following Lande (1976) and Turelli (1984). We have taken the mutation rate to be $u = 4 \times 10^{-4}$. Two values of the splitting of trait effects, Δ, have been used, namely $\Delta = m/5$ and $m/10$.

It is useful to have, for a general value of Ω, a quantitative measure of the degree of polymorphism. A choice that is suitable for our purposes is, for $\Omega = 1$, the ratio p_1/p_0, and generally is the ratio $p_{(1,0,0,\ldots,0)}/p_{(0,0,0,\ldots,0)}$. If this ratio is small compared with unity, the pattern of $p_{\mathbf{n}}$'s is sharply peaked and a low level of polymorphism exists, while if this ratio is close to unity, then there is a broad pattern of $p_{\mathbf{n}}$'s and a high level of polymorphism exists.

As an illustration of this measure of polymorphism for $\Omega = 1$, consider a low-mutation situation where we can apply an approximation scheme introduced by Turelli (1984) and related to a mutational scheme of Kingman (1978)—the "House of Cards" approximation. This applies when mutation is weak compared with selection: $u/(s\sigma^2) \ll 1$ and assuming that Δ is sufficiently small that $\Delta \ll u/(s\sigma)$, we find that to lowest non-trivial order in Δ, $p_1/p_0 \simeq 1 - 2\pi^{-1}s^2\sigma^2u^{-2}\Delta^2$. This ratio increases as Δ is decreased, and hence the level of polymorphism increases as Δ is decreased. See Fig. 1, where a numerically calculated example is illustrated. Details of the numerical method used for the production of the figures are given in Appendix B. While the method of Appendix B does not have large benefits for $\Omega = 1$, it offers a major reduction in the amount of computation required for higher values of Ω.

Since the ratio p_1/p_0 changes with Δ, it generally follows that quantities such as the genetic variance also change with Δ. Here we note that in the House of Cards regime $(u/(s\sigma^2) \ll 1)$, when Δ is small $(\Delta \ll u/(s\sigma))$, the genetic variance contains corrections to the $\Delta = 0$ continuum-of-alleles result. Estimates indicate that the correction terms are of order $\exp(-2\pi^2\kappa_1)$ or $\exp(-2\pi^2\sigma^2/\Delta^2)$ smaller than terms present in the $\Delta = 0$ result. For the range of parameters considered, $\kappa_1 \gg 1$ and $\sigma/\Delta \gg 1$, hence the Δ-dependent correction terms are extremely small. Thus, the genetic variance, in the parameter regime considered, is insensitive to the continuity or discreteness of allelic effects.

For $\Omega = 2$, the effective mutation rate, κ_2, is, by (11), independent of Δ. We do, however, see a broadening of the distribution of the $p_{\mathbf{n}}$'s, when Δ is decreased. See Fig. 2 for an example of this.

The effect seen follows since decreasing Δ, in $e^{-||\mathbf{n}-\mathbf{m}||^2\Delta^2/(2\sigma^2)}$, increases the range of trait categories accessible via a single mutation. To analytically illustrate what is occurring, consider a case where the "House of Cards" approximation is again applicable $(u/(s\sigma^2) \ll 1)$ and furthermore assume that Δ is sufficiently small that $\Delta \ll 2\sigma^2 \exp(-\pi^{-1}\kappa_2^{-1} - \gamma)$, where $\gamma \simeq 0.5772$ is Euler's constant. In this case, we find that to lowest non-trivial order in Δ, $p_{(1,0)}/p_{(0,0)} \simeq 1 - 2^{-1}\sigma^{-2}\Delta^2 \exp(2\pi^{-1}\kappa_2^{-1} + 2\gamma)$. This ratio, like the one for $\Omega = 1$, also increases as Δ decreases.

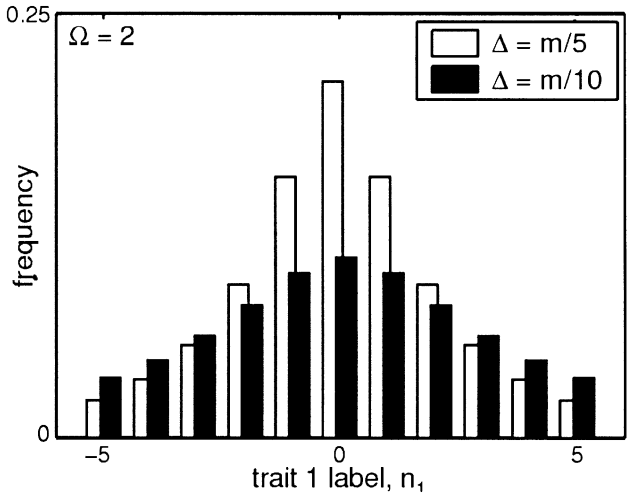

Fig. 2. For $\Omega = 2$, where each mutation affects two traits, the equilibrium frequencies of the different discrete categories of trait effects are $p_{\mathbf{n}} = p_{(n_1,n_2)}$. The marginal distribution describing trait 1 alone is $p^{(1)}_{n_1} = \sum_{n_2} p_{(n_1,n_2)}$ and this is plotted against the trait 1 label, n_1. Parameter values and details of the method used are the same as those of Fig. 1.

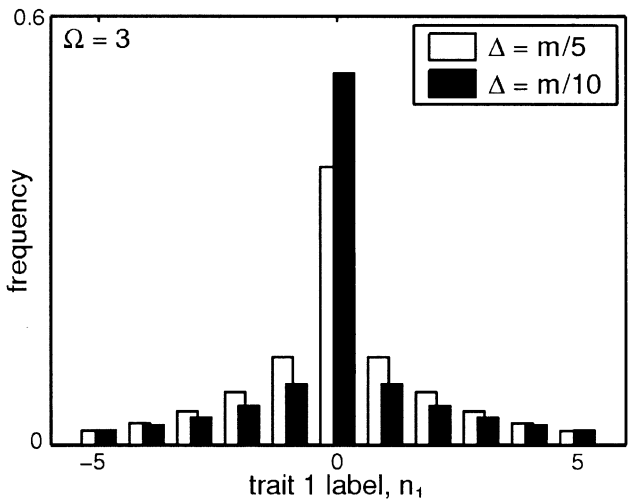

Fig. 3. For $\Omega = 3$, where each mutation affects three traits, the equilibrium frequencies of the different discrete categories of trait effects are $p_{\mathbf{n}} = p_{(n_1,n_2,n_3)}$. The marginal distribution describing trait 1 alone is $p^{(1)}_{n_1} = \sum_{n_2,n_3} p_{(n_1,n_2,n_3)}$ and this is plotted against the trait 1 label, n_1. Parameter values and details of the method used are the same as those of Fig. 1.

For $\Omega \geqslant 3$, we have $\kappa_\Omega \propto \Delta^{\Omega-2}$ and this decreases as Δ is decreased. If this were the only dependence on Δ, decreasing Δ would decrease the width of the distribution of $p_{\mathbf{n}}$'s. As discussed for the cases $\Omega = 1$ and $\Omega = 2$, the Δ dependence that resides in $e^{-||\mathbf{n}-\mathbf{m}||^2\Delta^2/(2\sigma^2)}$ has the tendency to broaden the distribution of the $p_{\mathbf{n}}$'s, when Δ is decreased. Thus, for $\Omega \geqslant 3$, the ultimate way the distribution of the $p_{\mathbf{n}}$'s changes, due to a decrease of Δ, is decided by which of the two opposing effects dominate. In Appendix C, we come to the non-rigorous conclusion that for sufficiently large Ω, the Δ dependence of κ_Ω will dominate the problem. Thus, beyond a critical Ω, termed Ω_c, a decrease in Δ causes a *decrease* in the width of the distribution of $p_{\mathbf{n}}$'s, i.e. corresponds to a reduction in the level of polymorphism. The value of Ω_c is determined by parameters in the problem; however, for all parameter choices, Ω_c can never be less than 3. We illustrate a case in Fig. 3 where $\Omega_c = 3$.

To analytically illustrate what is occurring, we again employ the "House of Cards" approximation (assuming $u/(s\sigma^2) \ll 1$) and furthermore assume Δ is sufficiently small that $\Delta^{\Omega-2} \ll u^{-1}s(2\pi\sigma^2)^{\Omega/2}$. We then find that to leading non-trivial order in Δ, $p_{(1,0,\ldots,0)}/p_{(0,0,\ldots,0)} \simeq \kappa_\Omega = us^{-1}(2\pi\sigma^2)^{-\Omega/2}\Delta^{\Omega-2}$. Thus, on decreasing Δ, the ratio $p_{(1,0,\ldots,0)}/p_{(0,0,\ldots,0)}$ decreases, signalling a narrower, less polymorphic distribution.

5. Discussion

In this work, the implications of discreteness of genotypic trait values, that followed from discreteness of the underlying allelic effects, has been investigated for models with different degrees of pleiotropy, Ω. In contrast to previous work involving discrete effects, we have adopted a Gaussian distribution of mutant effects. The parameter region investigated was effectively opposite that of stepwise models of mutation; thus, the model of mutation considered in this work had a large number of different trait values accessible from a single mutation.

Discreteness of trait values was characterised by the splitting of trait values, Δ, and it was shown that for mutations that affect only one or two characters, the effects of decreasing Δ increased the equilibrium polymorphism of the various categories of trait effects. For mutations that simultaneously affect Ω_c or more characters (where $\Omega_c \geqslant 3$), the effect of decreasing Δ is to decrease the equilibrium level of polymorphism of the various trait categories.

Given the results above, it is natural to ask how the behaviour seen in discrete-effect models can be reconciled with the singular behaviour of the distribution found in continuous-effect models for $\Omega \geqslant \Omega_c$. We can answer this question by considering the function that, when $\Delta \to 0$, goes over to the probability density describing a population of individuals with continuous trait values. This function is

$$\Phi_\Delta(\mathbf{x}) = \Delta^{-\Omega} \sum_{\mathbf{n}} \delta_{\mathbf{x},\mathbf{n}\Delta} p_{\mathbf{n}}, \tag{13}$$

where $\delta_{\mathbf{x},\mathbf{n}\Delta} = \prod_{i=1}^{\Omega} \delta_{x_i,n_i\Delta}$. The function $\Phi_\Delta(\mathbf{x})$ has the property that the mean value of any quantity that depends on genotypic trait values $\mathbf{x}$, say $A(\mathbf{x})$, is $\mathrm{A} = \Delta^{\Omega} \sum_{\mathbf{x}} \Phi_\Delta(\mathbf{x})A(\mathbf{x})$ and as $\Delta \to 0$ this becomes

$\int d^{\Omega}x\, \Phi(\mathbf{x})A(\mathbf{x})$ where $\Phi(\mathbf{x}) = \lim_{\Delta\to 0}\Phi_{\Delta}(\mathbf{x})$. Whether $\Phi(\mathbf{x})$ is finite (and non-singular) at $\mathbf{x}=0$ or has a singular behaviour (i.e. contains a Dirac delta function) at $\mathbf{x}=0$, depends crucially on how the central frequency, p_0, behaves when $\Delta\to 0$. It must be that for $\Omega<\Omega_c$, all frequencies including the central frequency, p_0, behave as a constant $\times\Delta^{\Omega}$, so that $\Phi_{\Delta}(\mathbf{x})$ remains finite and does not diverge as $\Delta\to 0$, see (13). Furthermore, for $\Omega\geqslant\Omega_c$, the central frequency, p_0, must remain finite, in order that a singular distribution is achieved. This follows since the contribution in (13) from $\mathbf{n}=\mathbf{0}$ is $\Delta^{-\Omega}\delta_{\mathbf{x},\mathbf{0}}p_{\mathbf{0}}$. If we let $C=\lim_{\Delta\to 0}p_{\mathbf{0}}$ and assume this is non-zero, the limiting form of $\Delta^{-\Omega}\delta_{\mathbf{x},\mathbf{0}}p_{\mathbf{0}}$ is a Dirac delta function, $\delta(\mathbf{x})$, multiplied by C.

Are these limiting behaviours consistent with the results we have presented so far for discrete effect alleles? Estimates for the case of small mutation rates $(u/(s\sigma^2)\ll 1)$, which lead to $\Omega_c=3$, and also allowed the "House of Cards" approximation to be employed, indicate that when Δ tends to zero, $p_{\mathbf{0}}\propto\Delta^{\Omega}$ for $\Omega=1$ and 2, while for $\Omega\geqslant 3$, it is found that $p_{\mathbf{0}}\propto\Delta^{0}$. In this way, we can reconcile the discrete and continuous-effect models, when $\Delta\to 0$.

As a final comment on this work, we note that the phenomena seen are not a result of the confounding issue that there are different strengths of selection, at different degrees of pleiotropy, since we have explicitly corrected for this. It follows that the results presented are a property of models incorporating pleiotropy.

Acknowledgments

It is a pleasure to thank John Welch for helpful suggestions and stimulating discussions. This research was supported by the Biotechnology and Biological Sciences Research Council (UK) under grant 85/G11043.

Appendix A

In this appendix, we consider the case where the splitting of trait values is large compared with the other scale in the problem, namely σ^2. This corresponds to $\Delta/\sigma\gg 1$ and when this holds we have, from (7) that $f_{\mathbf{nm}}\simeq(2\Omega)^{-1}$ if $||\mathbf{n}-\mathbf{m}||=1$ and $f_{\mathbf{nm}}$ is approximately zero for other values of $||\mathbf{n}-\mathbf{m}||$. This is the direct extension of the stepwise mutation model to higher than one dimension. For example, if $\Omega=2$ then $f_{\mathbf{nm}}\simeq(\delta_{n_1+1,m_1}\delta_{n_2,m_2}+\delta_{n_1-1,m_1}\delta_{n_2,m_2}+\delta_{n_1,m_1}\delta_{n_2+1,m_2}+\delta_{n_1,m_1}\delta_{n_2-1,m_2})/4$ where $\delta_{mn}=1$ if $m=n$ and is zero otherwise. We shall analyse the model for $\Omega=2$ with extensions to higher Ω straightforward.

When $\Delta/\sigma\to\infty$, we have (9) taking the form

$$(n_1^2+n_2^2)p_{(n_1,n_2)}-\frac{u}{4s\Delta^2}[p_{(n_1+1,n_2)}+p_{(n_1-1,n_2)}+p_{(n_1,n_2+1)}+p_{(n_1,n_2-1)}]=-\alpha p_{(n_1,n_2)}. \quad (A.1)$$

This equation is decomposable into two separate equations: set $p_{(n_1,n_2)}=q_{n_1}r_{n_2}$ and divide (A.1) by $q_{n_1}r_{n_2}$. This yields

$$n_1^2-\frac{u}{4s\Delta^2}\left[\frac{q_{n_1+1}}{q_{n_1}}+\frac{q_{n_1-1}}{q_{n_1}}\right]=-\left(n_2^2-\frac{u}{4s\Delta^2}\left[\frac{r_{n_2+1}}{r_{n_2}}+\frac{r_{n_2-1}}{r_{n_2}}\right]+\alpha\right)$$

and following standard arguments adopted for separable solutions of partial differential equations, we reason that the only way the left- and right-hand sides of the above equation can be equal for arbitrary choices of n_1 and n_2 is if they each equal a constant, say $-\alpha_1$. Thus, we have

$$n_1^2q_{n_1}-\frac{u}{4s\Delta^2}[q_{n_1+1}+q_{n_1-1}]=-\alpha_1q_{n_1},$$
$$n_2^2r_{n_2}-\frac{u}{4s\Delta^2}[r_{n_2+1}+r_{n_2-1}]=-\alpha_2r_{n_2}, \quad (A.2)$$

where $\alpha_2=\alpha-\alpha_1$. Both of the equations in (A.2) represent non-pleiotropic stepwise-mutation models and by symmetry, we expect $\alpha_1=\alpha_2$.

Thus, when $\Delta/\sigma\to\infty$, and we are in the pleiotropic stepwise-mutation regime, a full understanding of the distribution of different trait classes follows completely from knowledge of results for a non-pleiotropic stepwise mutation model. The generalisation to the case of arbitrary Ω is obvious. If q_n obeys $n^2q_n-\frac{u}{2\Omega s\Delta^2}[q_{n+1}+q_{n-1}]=-\alpha q_n$ with the q_n chosen so that $-\alpha$ has the smallest possible value, compatible with q_n non-negative and $\sum_n q_n=1$, then the distribution of the full (pleiotropic) problem is $p_{(n_1,n_2,\ldots,n_{\Omega})}=q_{n_1}q_{n_2}\ldots q_{n_{\Omega}}$.

Appendix B

In this appendix, details are given of the method used to numerically solve (10) for the equilibrium frequencies of trait categories.

Let $\{\mathbf{R}\}$ denote the complete set of different orthogonal matrices that preserve the length of $\mathbf{n}=(n_1,n_2,\ldots,n_{\Omega})^{\mathrm{T}}$, i.e. $||\mathbf{Rn}||^2=||\mathbf{n}||^2$, and have the effect that $\mathbf{Rn}$, like $\mathbf{n}$ itself, contains elements that are all signed integers. Thus, the effect of a typical $\mathbf{R}$ on $\mathbf{n}$ is to permute and change the sign of some elements; as a consequence, the only possible elements in each $\mathbf{R}$ are 0 or ± 1. By considering the possible effects of $\mathbf{R}$ on $\mathbf{n}$ we can infer that the total number of matrices in $\{\mathbf{R}\}$ is $2^{\Omega}\times\Omega!$.

To proceed, we replace $\mathbf{n}$ by $\mathbf{Rn}$ in (10) so it becomes $||\mathbf{n}||^2p_{\mathbf{Rn}}-\kappa_{\Omega}\sum_{\mathbf{m}(\neq\mathbf{Rn})}e^{-||\mathbf{Rn}-\mathbf{m}||^2\Delta^2/(2\sigma^2)}p_{\mathbf{m}}=-\alpha p_{\mathbf{Rn}}$. We

note that the sum over $\mathbf{m}$ can be replaced by a sum over $\mathbf{Rm}$ hence $||\mathbf{n}||^2 p_{\mathbf{Rn}} - \kappa_\Omega \sum_{\mathbf{Rm}(\neq \mathbf{Rn})} e^{-||\mathbf{Rn}-\mathbf{Rm}||^2 \Delta^2/(2\sigma^2)} p_{\mathbf{Rm}} = -\alpha p_{\mathbf{Rn}}$ and using orthogonality of $\mathbf{R}$ leads to $||\mathbf{n}||^2 p_{\mathbf{Rn}} - \kappa_\Omega \sum_{\mathbf{m}(\neq \mathbf{n})} e^{-||\mathbf{n}-\mathbf{m}||^2 \Delta^2/(2\sigma^2)} p_{\mathbf{Rm}} = -\alpha p_{\mathbf{Rn}}$ and this indicates that $p_{\mathbf{Rn}}$ obeys an identical equation to $p_{\mathbf{n}}$. Note that the $p_{\mathbf{n}}$ are restricted to be non-negative and sum to unity, and also that the eigenvector belonging to smallest eigenvalue, $-\alpha$, is, apart from a multiplicative factor, unique. As a result, there is a unique solution for the $p_{\mathbf{n}}$ and this indicates that $p_{\mathbf{Rn}} = p_{\mathbf{n}}$ for all admissible $\mathbf{R}$ and numerical investigations verify this. As a consequence, the $\mathbf{n}$'s, naturally form groups where the corresponding frequencies are identical. We shall call such groups of $\mathbf{n}$'s "classes," and as an example for $\Omega = 3$, consider the class that contains $\mathbf{n} = (1,0,0)^{\mathrm{T}}$ then other $\mathbf{n}$'s is the class are given by $(-1,0,0)^{\mathrm{T}}, (0,\pm 1,0)^{\mathrm{T}}$ and $(0,0,\pm 1)^{\mathrm{T}}$ making a total of six elements in this particular class (the largest possible number of elements in any class is $2^\Omega \times \Omega!$).

The above generalises the intuitive symmetries used by Slatkin (1987) to simplify some of his calculations.

The classes may be labelled by an integer $j = 0,1,2,\ldots$ and let us denote class j by C_j, and the corresponding value of $||\mathbf{n}||^2$, for all $\mathbf{n}$ in class j, is denoted by r_j. Let us introduce $\delta_{\mathbf{n},C_j}$ which has the value of 1 if $\mathbf{n} \in C_j$ and is zero otherwise. This satisfies $\sum_j \delta_{\mathbf{n},C_j} = 1$, since every $\mathbf{n}$ is a number of one class and the number of members of class j is

$$\Lambda_j \overset{\text{def}}{\equiv} \sum_{\mathbf{n}} \delta_{\mathbf{n},C_j}. \tag{B.1}$$

On multiplying (10) by $\delta_{\mathbf{n},C_j}$ and summing over $\mathbf{n}$ and also using $\sum_j \delta_{\mathbf{m},C_j} = 1$ leads to

$$(r_j + \kappa_\Omega) \sum_{\mathbf{n}} \delta_{\mathbf{n},C_j} p_{\mathbf{n}} - \kappa_\Omega \sum_k \sum_{\mathbf{n}} \sum_{\mathbf{m}} \delta_{\mathbf{n},C_j} \delta_{\mathbf{m},C_k} e^{-||\mathbf{n}-\mathbf{m}||^2 \Delta^2/(2\sigma^2)} p_{\mathbf{m}} = -\alpha \sum_{\mathbf{n}} \delta_{\mathbf{n},C_j} p_{\mathbf{n}}. \tag{B.2}$$

Let us define

$$\rho_j = \sqrt{\Lambda_j} p_{\mathbf{n}}, \quad \mathbf{n} \in C_j, \ j = 0,1,2,\ldots \tag{B.3}$$

which has the same value for all $\mathbf{n} \in C_j$ (the factor $\sqrt{\Lambda_j}$ has numerical advantages, as we shall shortly see). Then after some simplification, which includes using $\sum_{\mathbf{n}} \delta_{\mathbf{n},C_j} p_{\mathbf{n}} = \sqrt{\Lambda_j}\rho_j$, we can write (B.2) as

$$(r_j + \kappa_\Omega)\rho_j - \kappa_\Omega \sum_k \phi_{jk} \rho_k = -\alpha \rho_j, \tag{B.4}$$

where ϕ_{jk} is given by

$$\phi_{jk} = \frac{\sum_{\mathbf{n}} \sum_{\mathbf{m}} \delta_{\mathbf{n},C_j} \delta_{\mathbf{m},C_k} e^{-||\mathbf{n}-\mathbf{m}||^2 \Delta^2/(2\sigma^2)}}{\sqrt{\Lambda_j \Lambda_k}}. \tag{B.5}$$

If we had defined ρ_j, in (B.3), without the factor $\sqrt{\Lambda_j}$ then the matrix ϕ_{jk} would not have the numerical advantage of being symmetric. Note also, that the sum over $\mathbf{n}$ in ϕ_{jk} can effectively be omitted, since it can be proved that when $\mathbf{n}$ is a member of C_j, the sum $\sum_{\mathbf{m}} \delta_{\mathbf{m},C_k} e^{-||\mathbf{n}-\mathbf{m}||^2 \Delta^2/(2\sigma^2)}$ is independent of the particular $\mathbf{n}$ used. Because of this, we can write ϕ_{jk} in a form that is useful for computation:

$$\phi_{jk} = \sqrt{\frac{\Lambda_j}{\Lambda_k}} \sum_{\mathbf{m}} \delta_{\mathbf{m},C_k} e^{-||\mathbf{n}-\mathbf{m}||^2 \Delta^2/(2\sigma^2)}, \quad \mathbf{n} \in C_j. \tag{B.6}$$

A virtue of formulating the problem in terms of classes is that (B.4) is an equation for an object ρ_j with a single label, j, as opposed to the original representation of the eigenvalue equation, (10) which deals with the frequencies, $p_{\mathbf{n}}$, with Ω labels given by $n_1, n_2, \ldots, n_\Omega$, i.e. there is an effective dimensional reduction in the computational complexity.

Note that the normalisation condition on the $p_{\mathbf{n}}$, corresponds to the following normalisation on the ρ_j:

$$1 = \sum_{\mathbf{n}} p_{\mathbf{n}} = \sum_j \sum_{\mathbf{n}} \delta_{\mathbf{n},C_j} p_{\mathbf{n}} = \sum_j \Lambda_j \rho_j / \sqrt{\Lambda_j} = \sum_j \sqrt{\Lambda_j}\rho_j.$$

Note also that the marginal distribution of the different discrete categories of trait 1 effects is $p^{(1)}_{n_1} = \sum_{\mathbf{m}} \delta_{n_1,m_1} p_{\mathbf{m}}$ so e.g. for $\Omega = 3$, with $p_{\mathbf{n}} = p_{(n_1,n_2,n_3)}$, we have $p^{(1)}_{n_1} = \sum_{n_2,n_3} p_{(n_1,n_2,n_3)}$. We can write

$$p^{(1)}_{n_1} = \sum_j \sum_{\mathbf{m}} \delta_{n_1,m_1} \delta_{\mathbf{m},C_j} p_{\mathbf{m}} = \sum_j \sum_{\mathbf{m}} \delta_{n_1,m_1} \delta_{\mathbf{m},C_j} \frac{\rho_j}{\sqrt{\Lambda_j}}. \tag{B.7}$$

Truncating the calculations to a finite set of classes (say $j = 0,1,\ldots,J$) corresponds, in the full problem, involving the frequencies, $p_{\mathbf{n}}$, to a significantly larger number of frequencies than classes. Thus, working in terms of classes, which may be enumerated numerically, has considerable computational advantages and is the method used to produce the figures. Exact agreement is obtained for the special case $\Omega = 2$ with the predictions of the original equation, (10), without going through the transformations of this appendix.

Appendix C

In this appendix, we provide a suggestive, but non-rigorous analysis of what occurs to the distribution of the $p_{\mathbf{n}}$'s, when Δ is decreased. With $\beta = \alpha + \kappa_\Omega$, our starting point is (10), when written in the form $||\mathbf{n}||^2 p_{\mathbf{n}} - \kappa_\Omega \sum_{\mathbf{m}} e^{-||\mathbf{n}-\mathbf{m}||^2 \Delta^2/(2\sigma^2)} p_{\mathbf{m}} = -\beta p_{\mathbf{n}}$. This yields

$$p_{\mathbf{n}} = \frac{\kappa_\Omega \sum_{\mathbf{m}} e^{-||\mathbf{n}-\mathbf{m}||^2 \Delta^2/(2\sigma^2)} p_{\mathbf{m}}}{||\mathbf{n}||^2 + \beta}. \tag{C.1}$$

The ratio used in this work to characterise the level of polymorphism is

$$Q \overset{\text{def}}{\equiv} \frac{p_{(1,0,0,\ldots,0)}}{p_{(0,0,0,\ldots,0)}} \tag{C.2}$$

and using (C.1) we have

$$Q = \frac{\beta}{1+\beta} \frac{e^{-\Delta^2/(2\sigma^2)} \sum_{\mathbf{m}} e^{-2m_1\Delta^2/(2\sigma^2)} e^{-||\mathbf{m}||^2\Delta^2/(2\sigma^2)} p_{\mathbf{m}}}{\sum_{\mathbf{m}} e^{-||\mathbf{m}||^2\Delta^2/(2\sigma^2)} p_{\mathbf{m}}}$$

and for small Δ/σ we approximate Q by

$$Q \simeq \frac{\beta}{1+\beta}. \tag{C.3}$$

Since $\frac{\partial}{\partial\Delta}\frac{\beta}{1+\beta} = \frac{\partial\beta}{\partial\Delta}\frac{1}{(1+\beta)^2}$, it follows that the sign of $\partial\beta/\partial\Delta$ determines the behaviour of Q when Δ is decreased. In particular, if $\partial\beta/\partial\Delta > 0$ then Q will decrease when Δ is decreased. We can get a handle on $\partial\beta/\partial\Delta$ by employing the Hellman–Feynman theorem (Hellman, 1937; Feynman, 1939), which relates $\partial\beta/\partial\Delta$ to a derivative of the linear operator, of which β is the eigenvalue. For our purposes, this theorem takes the form

$$\frac{\partial\beta}{\partial\Delta} = \frac{\sum_{\mathbf{n},\mathbf{m}} p_{\mathbf{n}} \left[\frac{\partial}{\partial\Delta} H_{\mathbf{n},\mathbf{m}}\right] p_{\mathbf{m}}}{\sum_{\mathbf{n}} p_{\mathbf{n}}^2}, \tag{C.4}$$

where $H_{\mathbf{n},\mathbf{m}} = -||\mathbf{n}||^2\delta_{\mathbf{n},\mathbf{m}} + \kappa_\Omega e^{-||\mathbf{n}-\mathbf{m}||^2\Delta^2/(2\sigma^2)}$ and $\delta_{\mathbf{n},\mathbf{m}} = 1$ if $\mathbf{n} = \mathbf{m}$ and is zero otherwise. Using $\partial\kappa_\Omega/\partial\Delta = (\Omega-2)\kappa_\Omega/\Delta$ we find

$$\frac{\partial\beta}{\partial\Delta} = \frac{\kappa_\Omega}{\Delta} \frac{\sum_{\mathbf{n},\mathbf{m}} p_{\mathbf{n}} p_{\mathbf{m}} e^{-||\mathbf{n}-\mathbf{m}||^2\Delta^2/(2\sigma^2)} [(\Omega-2) - 2||\mathbf{n}-\mathbf{m}||^2 \frac{\Delta^2}{2\sigma^2}]}{\sum_{\mathbf{n}} p_{\mathbf{n}}^2}. \tag{C.5}$$

The origins of the terms in [] in this equation clearly show that when $\Omega \geqslant 3$, the Δ dependence of κ_Ω and $e^{-||\mathbf{n}-\mathbf{m}||^2\Delta^2/(2\sigma^2)}$ cause changes of opposite direction in β. Thus for $\Omega \geqslant 3$, the behaviour of β (and hence the level of polymorphism) that follows when Δ is decreased, depends upon which of the two opposing effects dominate.

Let us consider, for $\Omega \geqslant 3$, the condition under which $\partial\beta/\partial\Delta > 0$ holds, or equivalently when the numerator of the right-hand side of (C.5) is positive. This condition can be written as

$$(\Omega-2) + 2\frac{\partial}{\partial\lambda} \ln \left[\sum_{\mathbf{n},\mathbf{m}} p_{\mathbf{n}} p_{\mathbf{m}} e^{-\lambda||\mathbf{n}-\mathbf{m}||^2\Delta^2/(2\sigma^2)} \right] \Bigg|_{\lambda=1} > 0. \tag{C.6}$$

There are two straightforward things we can learn from this relation:

(i) If $\sum_{\mathbf{n}} p_{\mathbf{n}} ||\mathbf{n}||^2 \Delta^2/\sigma^2 \equiv \mathrm{Var}(\mathbf{n})\Delta^2/\sigma^2 \ll 1$ then the condition $\partial\beta/\partial\Delta > 0$ can be approximately written as $(\Omega-2) - 2\,\mathrm{Var}(\mathbf{n})\Delta^2/\sigma^2 > 0$ and because we have assumed a small variance, this automatically leads to $\partial\beta/\partial\Delta > 0$ for $\Omega > 2$ i.e. $\Omega_c = 3$.

(ii) If we assume $p_{\mathbf{n}}$ is a Gaussian: $p_{\mathbf{n}} \propto \exp[-||\mathbf{n}||^2\Delta^2/(2a^2)]$ for some a, then for small Δ, we can approximate the sums in (C.6) by integrals. A short calculation indicates that $\partial\beta/\partial\Delta > 0$ leads to $\Omega/(1+2a^2/\sigma^2) - 2 > 0$. Thus for any value of a, we can always find an Ω for which the inequality is satisfied. This makes it plausible that in a more general case, where there is a unimodal distribution of $p_{\mathbf{n}}$'s of arbitrary width (corresponding to the value of a in the above example), it is always possible to find a sufficiently large Ω, such that decreasing Δ results in the influence of κ_Ω dominating the influence of $e^{-\lambda||\mathbf{n}-\mathbf{m}||^2\Delta^2/(2\sigma^2)}$, and hence decreasing Δ results in a decrease in the level of polymorphism.

References

Apostal, T.M., 1979. Mathematical Analysis, 2nd Edition. Addison–Wesley, Reading, MA.

Barton, N.H., 1990. Pleiotropic models of quantitative variation. Genetics 124, 773–782.

Barton, N.H., 1998. Evolutionary biology—the geometry of adaptation. Nature 395, 751–752.

Barton, N.H., Keightley, P.D., 2002. Understanding quantitative genetic variation. Nat. Rev. Genet. 3, 11–21.

Barton, N.H., Turelli, M., 1989. Evolutionary quantitative genetics: how little do we know? Ann. Rev. Genet. 23, 337–370.

Caspari, E., 1952. Pleiotropic gene action. Evolution 6, 1–18.

Crow, J.F., Kimura, M., 1964. The theory of genetic loads. Proceedings of the XIth International Congress of Genetics 2, 495–505.

Feynman, R.P., 1939. Forces in molecules. Phys. Rev. 56, 340–343.

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

Hellman, H., 1937. Einfuhrung in die Quantumchemie. Franz Deuticke, Leipzig.

Keightley, P.D., Hill, W.G., 1990. Variation maintained in quantitative traits with mutation–selection balance: pleiotropic side-effects on fitness traits. Proc. R. Soc. London B 242, 95–100.

Kimura, M., 1965. A stochastic model concerning the maintenance of genetic variability in quantitative characters. Proc. Natl. Acad. Sci. USA 54, 731–736.

Kingman, J.F.C., 1978. A simple model for the balance between selection and mutation. J. Appl. Probab. 15, 1–12.

Lande, R., 1976. The maintenance of genetic variability by mutation in a polygenic character with linked loci. Genet. Res. 26, 221–235.

Lande, R., 1980. The genetic covariance between characters maintained by pleiotropic mutation. Genetics 94, 203–215.

Lyman, R.F., Mackay, T.F.C., 1998. Candidate quantitative trait loci and naturally occurring variation for bristle number in *Drosophila melanogaster*: the Delta–Hairless gene region. Genetics 149, 983–998.

Nuzhdin, S.V., Dilda, C.L., Mackay, T.F.C., 1999. The genetic architecture of selection response: inferences from fine-scale mapping of bristle number quantitative trait loci in *Drosophila melanogaster*. Genetics 153, 1317–1331.

Orr, H.A., 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution 52, 935–949.

Orr, H.A., 2000. Adaptation and the cost of complexity. Evolution 54, 13–20.

Robertson, A., 1967. The nature of quantitative variation. In: Brink, A. (Ed.), Heritage from Mendel. The University of Wisconsin Press, Madison, WI, pp. 265–280.

Slatkin, M., 1987. Heritable variation and heterozygosity under a balance between mutations and stabilizing selection. Genet. Res. 50, 53–62.

Slatkin, M., Frank, S., 1990. The quantitative genetic consequences of pleiotropy under stabilizing and directional selection. Genetics 125, 207–213.

Turelli, M., 1984. Heritable genetic variation via mutation selection balance: Lerche's zeta meets the abdominal bristle. Theor. Popul. Biol. 25, 138–193.

Turelli, M., 1985. Effects of pleiotropy on predictions concerning mutation–selection balance for polygenic traits. Genetics 111, 165–195.

Wright, S., 1968. Evolution and the Genetics of Populations. University of Chicago Press, Chicago.

Wagner, G.P., 1989. Multivariate mutation selection balance with constrained pleiotropic effects. Genetics 122, 223–234.

Waxman, D., Peck, J.R., 1998. Pleiotropy and the preservation of perfection. Science 279, 1210–1213.

Waxman, D., Peck, J.R., 2000. The outcome of evolution when mutations are highly pleiotropic. Selection 1, 181–191.