Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Theoretical Biology 239 (2006) 152–160
Journal of Theoretical Biology
www.elsevier.com/locate/yjtbi

Near-periodic substitution and the genetic variance induced by environmental change

Y. Bello, D. Waxman*

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

Received 11 February 2005; received in revised form 27 April 2005; accepted 28 April 2005
Available online 2 November 2005

Abstract

We investigate a model that describes the evolution of a diploid sexual population in a changing environment. Individuals have discrete generations and are subject to selection on the phenotypic value of a quantitative trait, which is controlled by a finite number of bialleic loci. Environmental change is taken to lead to a uniformly changing optimal phenotypic value. The population continually adapts to the changing environment, by allelic substitution, at the loci controlling the trait. We investigate the detailed interrelation between the process of allelic substitution and the adaptation and variation of the population, via infinite population calculations and finite population simulations. We find a simple relation between the substitution rate and the rate of change of the optimal phenotypic value.

Keywords: Quantitative trait; Substitution; Genetic variance; Stabilizing selection; Environmental change

1. Introduction

The evolution of populations rarely, if ever, takes place in a static environment. Apart from purely abiotic changes of a physical environment, there may also be changes due to the interaction of a population with other populations (Van Valen, 1973; or, for recent developments, see Gavrilets, 1997). Evolution is thus inexorably influenced by a changing environment and this may have implications for a variety of subjects including the evolution of sex (Maynard Smith, 1978; Waxman and Peck, 1999; Bürger, 1999). In the present work, we consider a population characterized by a single quantitative trait that possesses an optimal phenotypic value, because it is subject to stabilizing selection. Following previous initiatives, we model environmental change by a constant rate of change of the optimal phenotypic value (see e.g. Charlesworth, 1993; Bürger and Lynch, 1995; Waxman and Peck, 1999; Bürger, 1999). The present work has the closest relation with the work of Waxman and Peck (1999), which dealt with a very large (effectively infinite) population of individuals. In that work it was found that a steady-state situation became established, where the population tracks (with a lag) the changing environment. It was also found that there were extremely large enhancements in the genetic variance associated with very modest rates of environmental change. Indeed, from the first two columns of Table 1 of the paper by Waxman and Peck (1999), which applies for 10 diploid loci, it may be inferred that in a sexual population, changing the optimal phenotypic value by a small amount, e.g. 0.01% or 0.1% of an environmental standard deviation, each generation, leads to the genetic variance being increased to 450% or 1400% of its value in a static environment. These very large increases indicate a significant sensitivity (or lack of robustness) of the genetic variance to a changing environment. Indeed, because of this, we can conclude that the knowledge of just the strength of selection and the size of mutation rates may not be sufficient to predict the level of genetic variance of a population.

The increase in the genetic variance, referred to above, must, ultimately, originate from processes of allelic

*Corresponding author. Tel.: +44 1 273 678 559; fax: +44 1 273 678 433.
E-mail address: D.Waxman@Sussex.ac.uk (D. Waxman).

0022-5193/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2005.08.044

Table 1
Comparison of the prediction, that the mean time interval between substitutions is $\rho^{-1} = 2m/\alpha$, (as follows from Eq. (6)), with the results of numerical simulation

α	$\rho^{-1} = \frac{2m}{\alpha}$	Time interval (simulation): Mean ± standard deviation
0.5×10^{-4}	8.0×10^3	$(7.7 \pm 2.3) \times 10^3$
1.0×10^{-4}	4.0×10^3	$(4.2 \pm 1.3) \times 10^3$
2.0×10^{-4}	2.0×10^3	$(1.9 \pm 1.6) \times 10^3$
3.0×10^{-4}	1.3×10^3	$(1.2 \pm 0.6) \times 10^3$

The quantity α is the rate of change of the optimal phenotypic value and m is the scale of allelic effects of all loci. In the simulations, the population size was 10 000 individuals, with each adult producing a mean number of 1.5 offspring each generation. Parameter values adopted were: $u = 10^{-5}$, $m = 0.2$, $-|Z_{opt,0}| = -3$ and $V_s = 21$ (see text for a description of these).

substitution, as the population continually adapts to the changing environment. The work of Waxman and Peck (1999) was made in the framework of a continuum of alleles model (Crow and Kimura, 1964; Kimura, 1965), which assumes a very large number of alleles at any locus. In the present work, we investigate the detailed interrelation between the process of allelic substitution and the adaptation and variation of a population in a changing environment. Accordingly, we consider a model where the substitution process is more transparent than one with many alleles available at each locus, and the simplest and most transparent of such models has only two alleles at any locus. As we shall see, such work may have relevance to recent work on genetic variation in *Drosophila melanogaster* by Gardner et al. (2004).

2. Model

We consider a panmictic population of sexual organisms that are diploid and dioecious. The lifecycle of the population, that takes place in discrete generations, is: (i) random union of gametes to form zygotes, (ii) maturation to adulthood, with viability selection, (iii) production of gametes by Mendelian segregation; mutation is taken to occur during gamete production, (iv) death of adults. Census is made at the zygotic stage and at this time, the population is in Hardy–Weinberg equilibrium.

Individuals do not exhibit any sexual dimorphism, and are characterized by a single phenotypic trait that is additively controlled by the effects of $2n$ alleles at n unlinked loci. At locus $j(= 1, 2, \ldots, n)$ we take there to be only two possible alleles. We label these as B_j and b_j, and they, respectively, contribute $m_j/2$ and $-m_j/2$ to the value of the trait, where $m_j > 0$. We shall sometimes refer to m_j as the *scale* of allelic effects of locus j. The phenotypic value of the trait, Z, consists of a sum of its genotypic value, G, and a statistically independent random environmental effect ε, thus $Z = G + \varepsilon$. An individuals genotypic value is given by

$$G = \sum_{j=1}^{n} m_j(x_j + y_j)/2, \tag{1}$$

where $x_j(y_j)$ is a variable indicating the state of locus j of maternal (paternal) origin and only takes the values ± 1. Thus G is restricted to the range $\sum_{j=1}^{n} m_j \geqslant G \geqslant -\sum_{j=1}^{n} m_j$. The random environmental effect, ε, is normally distributed with mean zero. Following convention, an overall scale of units for various quantities is chosen, so that ε has a variance unity.

The values of x_j and y_j are, for all j, assumed to be identical to the parental values unless a mutation occurs in the production of gametes. We assume mutations occur independently to different alleles and that the rate (i.e. probability) of mutation at locus j, between B_j and b_j in either direction, is u_j.

Fitness is taken to be determined entirely by Gaussian stabilizing viability selection on the phenotypic value of the trait. The relative fitness of individuals of *genotypic* value G arises from an average of viability over environmental effects (see e.g. Turelli, 1984 or Bulmer, 1989) and is given by

$$w(G) = \exp[-(G - Z_{opt})^2/(2V_s)], \tag{2}$$

where $V_s^{-1}(> 0)$ is a direct measure of the intensity of selection on genotypic values of the trait and Z_{opt} is the optimal phenotypic value (and also the optimal genotypic value).

In what follows, we shall assume weak selection ($V_s^{-1} \ll 1$), as is often observed in naturally occurring populations (Turelli, 1984). Identical or very closely related models have been studied by Wright (1935), Barton (1986), Maynard Smith (1988), Bulmer (1989) and a number of other authors.

3. Change in the optimal phenotypic value

Let us return to the properties of selection. As we have said in the Introduction, the optimal phenotypic value, Z_{opt}, may be influenced by interactions of the population in question, with other populations or by the physical environment. As a consequence, Z_{opt} generally depends on time, t, and we take

$$Z_{opt}(t) = \begin{cases} -|Z_{opt,0}|, & t \leqslant 0, \\ -|Z_{opt,0}| + \alpha t, & t > 0. \end{cases} \tag{3}$$

This corresponds to the optimal phenotypic value having a fixed negative value of $-|Z_{opt,0}|$ for times $t \leqslant 0$ (i.e. corresponding to a static environment), while for times $t > 0$, the optimal phenotypic value changes at a constant rate of α per generation. By virtue of the choice of units adopted (that ensure the variance of the environmental effects is unity) the quantity α represents the change in the optimal phenotypic value, in units of the *standard deviation of environmental effects*.

Analytical studies of related asexual models, where there are many discrete effect alleles possible at each locus are given in the work of Broom et al. (2003).

4. Results for an infinite population

We first consider a very large, effectively infinite population, where effects of random genetic drift are negligible. We neglect linkage disequilibria. Investigations of the full multilocus problem in static environments indicate that neglect of linkage disequilibria is a very reasonable approximation when selection is weak (see e.g. Bulmer, 1989; Turelli and Barton, 1990). In a changing environment, this neglect of linkage disequilibria requires some discussion. The analysis presented by Bulmer (1974) for the infinitesimal model (see also Waxman, 2000) indicates that the chief effect of linkage disequilibria, when selection is stabilizing, is the production of negative correlations that reduce the genetic variance. The fractional reduction in genetic variance is proportional to V_G/V_s where V_G is the instantaneous genetic variance (strictly, we should use the genic variance—the genetic variance when calculated assuming linkage equilibrium) and V_s^{-1} is the intensity of selection on genotypic values (V_s^{-1} appears in Eq. (2)). The values we find for V_G, below, indicate that linkage disequilibria have only a very small effect on the genetic variance, of order 1%, and hence can reasonably be neglected. Bulmer's result is likely to be an overestimate of the effect of linkage disequilibria, when the optimum is moving, since the lag of the mean trait value, behind the optimal trait value, means selection has a directional aspect to it. Indeed, if fitness can be approximated as an exponential function of trait values, then such a function is multiplicative across loci, and does not produce any linkage disequilibria.

A consequence of Hardy–Weinberg equilibrium and the neglect of linkage disequilibria, is the statistical independence of all alleles both across and between loci.

Let $p_j(q_j \equiv 1-p_j)$ denote the frequency of the allele $B_j(b_j)$ of maternal origin at locus j in a particular generation and $p_j'(q_j')$ the corresponding frequency in the following generation. Apart, possibly, from the initial generation, the frequency of paternal origin alleles coincides with that of maternal origin and we shall henceforth assume this.

Given that the intensity of selection, V_s^{-1}, is small, the analysis presented in e.g. Bulmer (1989) then applies. Let $\mathbf{p}=(p_1,p_2,\ldots,p_n)$ denote the set of frequencies of the B allele at the different loci. The dynamical equations determining the change of allele frequencies can be written for $j=1,2,\ldots,n$ as

$$p_j' = p_j + \frac{1}{2}p_j q_j \frac{\partial}{\partial p_j}\Phi(\mathbf{p}), \tag{4}$$

where, equivalent to Eqs. (9) and (13) of Bulmer (1989), we may take

$$\Phi(\mathbf{p}) = \ln(E[w(G)]) + \sum_{j=1}^{n} 2u_j \ln[p_j(1-p_j)]. \tag{5}$$

The first term in $\Phi(\mathbf{p})$ arises from selection and involves $E[w(G)]$, which denotes the mean (or expected value) of relative fitness, while the second term arises from mutation.

With $E[G]$ and V_G denoting the mean genotypic effect and the genetic variance, we establish, in Appendix A, that when $V_G/V_s \ll 1$ and $(E[G]-Z_{opt})^2/V_s \ll 1$, it is a good approximation to replace $\ln(E[w(G)])$ in Eq. (5) by $E[\ln(w(G))]$ (which is given in Eq. (11) of Appendix A).

Since $V_G \leqslant n\overline{m^2}/2$, where $\overline{m^2}=\sum_{j=1}^{n} m_j^2/n$ is the mean square scale of allelic effects, the condition $V_G/V_s \ll 1$ can be expressed as $n\overline{m^2}/(2V_s) \ll 1$. As typical parameter-values, we take $n=10$, $\sqrt{\overline{m^2}}=0.2$, $u_j \sim 10^{-5}$ and $V_s=20$ (Turelli, 1984; see also Lynch and Walsh, 1998, Chapter 12) so $V_G/V_s \ll 1$ is well satisfied. The value of $\sqrt{\overline{m^2}}$ adopted corresponds to a standard deviation of mutant effects of order 0.2. Additionally, for the parameter values adopted we have found that virtually all variation in allele frequencies (the p_j) occurs where $(E[G]-Z_{opt})^2/V_s \ll 1$ and hence using $E[\ln(w(G))]$ in place of $\ln(E[w(G)])$ is justified. The explicit form of Eq. (4) may be found by carrying out the differentiations in this equation and coincides with the sum of Eqs. (12) and (13) of Bulmer (1989).

When Z_{opt} has a constant value, Barton (1986) has found that there are multiple equilibrium solutions of Eq. (4) and that these have different genetic variances associated with them. This indicates that the interplay between selection and mutation, in this biallelic system, results in multiple equilibria; there is more than one local maximum of $\Phi(\mathbf{p})$ of Eq. (5). The same nonlinear forces manifest themselves when Z_{opt} changes with time, according to Eq. (3), leading to a complex dynamical behaviour of the allele frequencies (the p_j).

It is tempting, as a first approach, to consider the case of *equivalent loci* (also called interchangeable loci), where the scale of allelic effects at all loci are identical: $m_j=m$ and where there are the same allelic mutation rates at all loci: $u_j=u$. For this case, we iterated the dynamical equation, Eq. (4) from a large negative time, $-T$, and took $-|Z_{opt,0}|$ sufficiently negative that up to time $t=0$, only b alleles were selectively favourable at every locus. Provided $-T$ was sufficiently large and negative, it was found that irrespective of the initial state of the population, all loci were essentially fixed at the b allele (i.e. all $q_j \simeq 1$), by the time $t=0$ was reached. As a consequence of this and the equivalence of all loci, all allele frequencies were completely synchronized for times $t>0$. The result for the allele frequencies is displayed in Fig. 1.

We view this as a *spurious synchronization* of allelic effects that is not of biological interest. It arises from the extreme symmetry of the problem: the complete

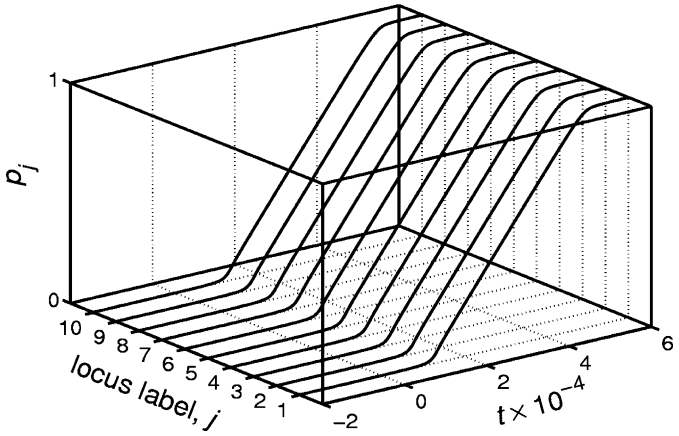

Fig. 1. The frequencies p_j of the B alleles at the 10 different loci are plotted against time t. The figure was produced by numerically iterating the dynamical equation, Eq. (4). The parameter values adopted were: $u = 10^{-5}$, $\alpha = 10^{-4}$, $-|Z_{opt,0}| = -3$ and $V_s = 21$. The scale of allelic effects were taken as $m_j = 0.2$, for all j. The results displayed are for the allele frequencies after transients have died away and so for the initial time shown, $t = -2 \times 10^4$, all p_j were close to 0. All allele frequencies are synchronized.

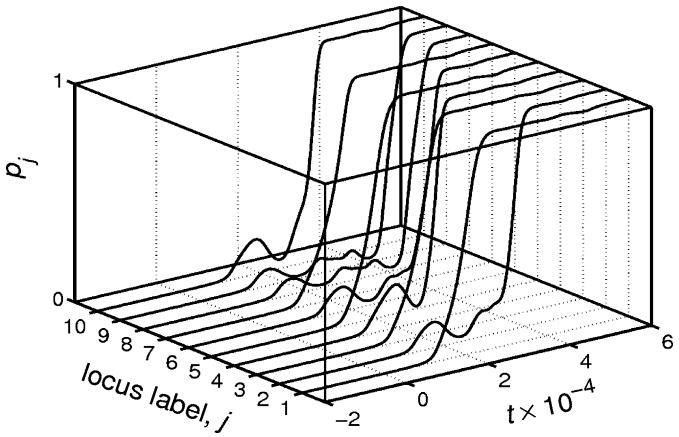

Fig. 2. The frequencies p_j of the B alleles at the 10 different loci are plotted against time t. The figure was produced by numerically iterating the dynamical equation, Eq. (4). The parameter values adopted were: $u = 10^{-5}$, $\alpha = 10^{-4}$, $-|Z_{opt,0}| = -3$ and $V_s = 21$. The scale of allelic effects at different loci (the m_j), were independently drawn at random from uniformly distributed numbers over $[0.15, 0.25]$. The results displayed are for the allele frequencies after transients have died away. For the earliest time shown, $t = -2 \times 10^4$, all p_j were close to 0.

equivalence of all n loci and the fact that all B alleles had identical frequencies at time $t = 0$. Indeed, such a model could be analysed in the framework of the hypergeometric model (Barton, 1992; Shpak and Kondrashov, 1999; Barton and Shpak, 2000), where the solutions are unstable for many forms of selection. There are other cases in the literature where equivalence of loci leads to unrepresentative effects (Welch and Waxman, 2002; Waxman and Welch, 2003; Waxman and Peck, 2003). Let us therefore consider a more realistic situation where there is some variation of the scale of allelic effects across loci. We have thus drawn the m_j at random from a uniform distribution centred on 0.2 (see Fig. 2 for further details). This yields a very different pattern of changing allele frequencies, as seen in Fig. 2, indicating that the original pattern in Fig. 1 is unstable to small deviations from equivalence of loci. In Figs. 3 and 4 we plot the corresponding mean genotypic value and genetic variance against time.

The behaviour exhibited in Figs. 2–4 are complicated, nevertheless, some general features are apparent, or can be inferred from these figures or the underlying data.

1. Different loci generally undergo substitutions at different times (Fig. 2).

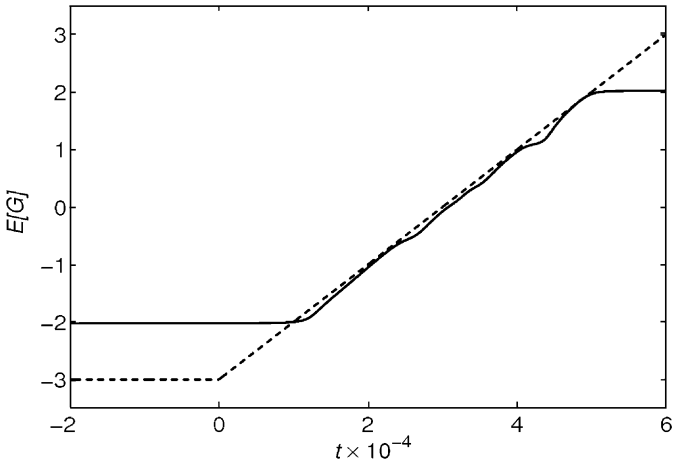

Fig. 3. The mean genotypic effect, $E[G]$, is plotted against time (solid line). The figure was produced by numerically iterating the dynamical equation, Eq. (4) and corresponds to the allele frequencies of Fig. 2. Parameter values are those given in Fig. 2. Also plotted in the same figure is the optimal phenotypic value, Z_{opt} (dashed line). The genotypic values of the trait range from approximately -2 to 2. When the optimal phenotypic value lies within this range, the population tracks the changing optimum, with a small but non-zero lag: $E[G] < Z_{opt}$.

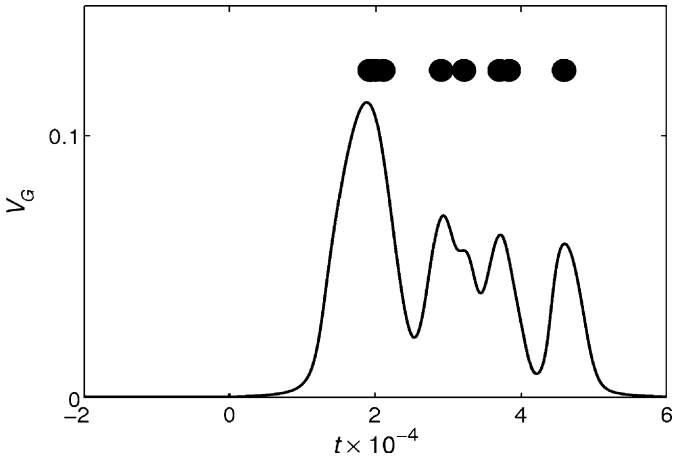

Fig. 4. The genetic variance, V_G, is plotted against time. The figure was produced by numerically iterating the dynamical equation, Eq. (4), and corresponds to the allele frequencies of Fig. 2. Parameter values are those given in Fig. 2. The horizontal row of black dots represent the times at which different substitutions occurred. There should be 10 dots visible, corresponding to substitutions at the 10 loci, however some of the dots overlap, since some substitutions occurred within a very short time of one another.

2. The occurrence of substitution at locus j is taken as the time when the frequency of the B_j allele achieves $p_j = 1/2$. Except when substitutions of some loci occur very close together, there appears to be a simple pattern connecting the size of m_j and the substitution order. The locus with smallest m_j undergoes substitution first, the locus with next largest m_j undergoes substitution next.
3. The substitutions allow the mean genotypic value, $E[G]$, to follow the changing optimum, albeit with a slightly variable lag (Fig. 3.). This process can continue until all genetic variation is exhausted.
4. Genetic variation is exhausted at a time slightly larger (due to the lag) than the time $t_{\max}$ given by $-|Z_{opt,0}| + \alpha t_{\max} = \sum_{j=1}^{n} m_j$. For times appreciably larger than $t_{\max}$ the population will go extinct if the environmental change persists.
5. The contribution to the genetic variance from locus j is $2m_j^2 p_j(1-p_j)$ and has a maximum value of $m_j^2/2$, at $p_j = 1/2$ (the value of p_j corresponding to maximum polymorphism). Thus the peaks in the genetic variance that are evident in Fig. 4, for $t>0$, result from the frequency of the B alleles passing through the value of $p = 1/2$. When substitutions occur close together, the resulting features in the genetic variance are, because of linkage equilibrium, simply the sum of contributions from different loci. Because of this additive property, the height of peaks in the genetic variance can be appreciably larger than $m_j^2/2$ if a number of substitutions contribute to the same peak.

5. Results for a finite population

In addition to numerical iteration of the dynamical equation, Eq. (4), we have performed numerical simulations of finite populations. In contrast to the infinite population calculations of the previous section, numerical simulations do not neglect linkage disequilibria. Thus, in the simulations, linkage disequilibria are fully incorporated into the dynamics.

The lifecycle is as outlined, in Section 2, with each individual producing an average of 1.5 offspring each generation and a fixed number of $N = 10\,000$ adults were maintained, each generation (by non-selectively thinning the population, after stage (ii) of the lifecycle). We have found it makes little difference whether the scale of allelic effects (the m_j) are the same for every locus or whether we use the range used for Figs. 2–4. Accordingly, we have taken all m_j to have the same value: $m_j = m = 0.2$.

It is possible to provide an estimate of the expected time interval between substitutions. We have specified the *time* of a substitution, at locus j, as that time at which $p_j = 1/2$. We then reason that as a result of a substitution

$$\begin{pmatrix}\text{change in optimal}\\ \text{phenotypic value}\end{pmatrix} = \begin{pmatrix}\text{change of effect of the locus}\\ \text{that underwent the substitution}\end{pmatrix}.$$

With ρ denoting the mean rate of substitutions and ρ^{-1} denoting the mean time between substitutions, we have that the change in optimal phenotypic value is (rate of change of optimum)×(time between substitutions), which is $\alpha \times \rho^{-1}$. We also note that substitution of both b alleles (each with effect $-m/2$), at a locus, by two B alleles (each with effect $+m/2$) results in a net change of effect on the trait of $2m$. Equating the change in optimal phenotypic value to the change in effect of the trait: $\alpha \times \rho^{-1} = 2m$, leads to a mean rate of substitutions of

$$\rho = \frac{\alpha}{2m}. \tag{6}$$

Thus ρ is proportional to the rate of change of the optimal phenotypic value.

We have performed numerical simulations for a population size of 10 000 individuals and explored the behaviour of the population when the rate of change of the optimal phenotypic value takes the values $\alpha = 0.5\times10^{-4}$, 1×10^{-4}, 2×10^{-4} and 3×10^{-4}.

In Fig. 5 we plot the changing allele frequencies, as a function of time, when the rate of change of the optimum is $\alpha = 10^{-4}$. In Figs. 6 and 7 we plot the corresponding mean genotypic value and genetic variance as functions of time.

It is clear, from Fig. 6 that the mean genotypic value closely tracks the moving optimum, as it did when the population was effectively infinite.

By contrast to the infinite population results, it is clearly seen, in Fig. 7, that the substitutions occur in a *very* regular, near-periodic manner. For $n = 10$ loci, there are 10 substitutions possible (assuming all $p_j \simeq 0$ at time $t = 0$; as occurs if $-|Z_{opt,0}|$ is sufficiently negative). For a time interval a little larger than $10\rho^{-1}$ all substitutions occur. This is very different behaviour to a Poisson point process, which is *substantially* noisier. A Poisson point process, over

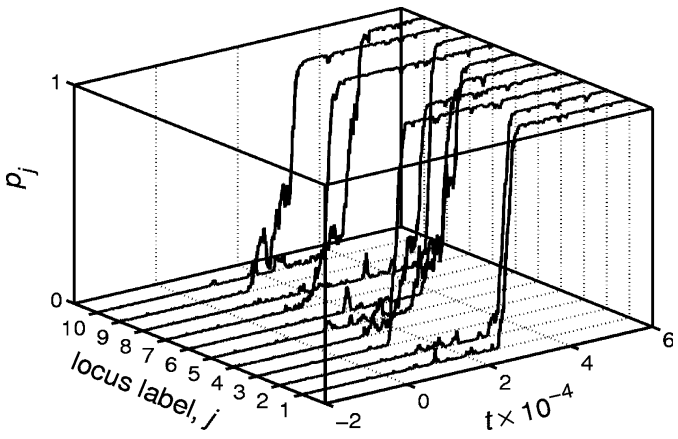

Fig. 5. The frequencies p_j of the B alleles at the 10 different loci are plotted against time t. The figure was produced by numerically simulating a population of $N = 10\,000$ individuals, with each adult producing a mean number of 1.5 offspring each generation. The other parameter values adopted were: $u = 10^{-5}$, $\alpha = 10^{-4}$, $-|Z_{opt,0}| = -3$ and $V_s = 21$. The scale of allelic effects at all loci were taken as $m_j = 0.2$. The results displayed are for the allele frequencies after transients have died away. For the earliest time shown, $t = -2\times10^4$, all p_j were close to 0.

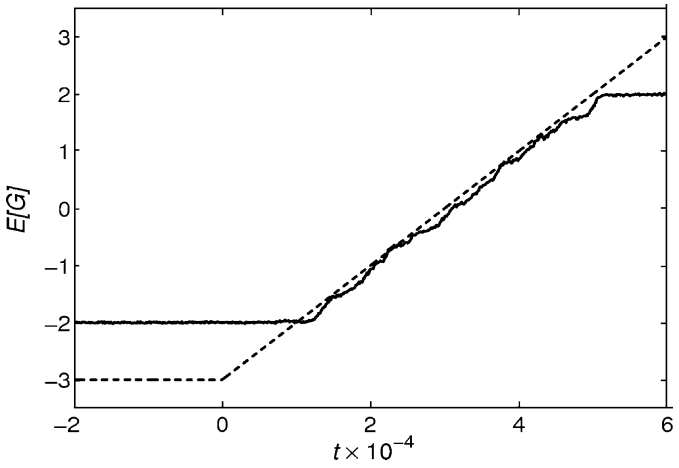

Fig. 6. The mean genotypic effect, $E[G]$, is plotted against time (solid line). The figure was produced by numerically simulating a population, as described in Fig. 5. Also plotted in the same figure is the optimal phenotypic value, Z_{opt} (dashed line).

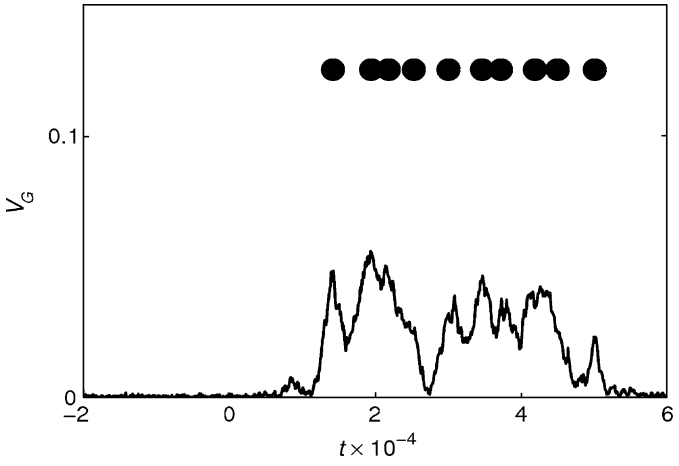

Fig. 7. The genetic variance, V_G, is plotted against time. The figure was produced by numerically simulating a population, as described in Fig. 5. There is very close agreement with the genetic variance calculated from the allele frequencies at different loci, assuming linkage equilibrium, $V_G = 2m^2 \sum_{j=1}^{n} p_j(1-p_j)$. The horizontal row of black dots represent the times at which different substitutions occurred.

any time interval where the expected number of substitutions is k, would lead to the order of $k \pm \sqrt{k}$ substitutions (since the variance in the number of substitutions equals the expected number itself, in such a process).

We can provide a quantitative measure of how regularly the substitutions occur by considering the index of dispersion of the substitution process, R, which, for a given time interval, is the ratio of the variance in the number of substitutions that occur, to the mean number of substitutions. It seems most relevant to determine R over a time interval where the population is closely tracking the changing optimum. Accordingly, we have carried out 20 independent simulations and determined R for the time interval ranging from 2×10^4 generations to 4×10^4 generations; from Figs. 3 or 6, this time interval is wholly contained in the time region where the population closely tracks the changing optimum. A Poisson process would lead to a value of R of unity; whereas we find values of R that are substantially smaller than unity, thereby indicating that the random process underlying the present model is substantially more periodic (or less random) than a Poisson process. In particular, when all of the m_j's have identical values ($m_j = 0.2$) or when the m_j's differ from locus to locus, (we used the set $0.1693, 0.2182, 0.1803, 0.2042, 0.1651, 0.2198, 0.1878, 0.2360, 0.2354, 0.2094$), we find values for the index of dispersion, R, that are smaller than 0.1.

The simulations used to determine R (where the m_j's differed from locus to locus) also allowed us to determine if there is any pattern in the order of substitutions at different loci and the sizes of the m_j. For example, is the locus with the largest (or smallest) value of m_j the first (or last) locus to undergo a substitution in a finite population? We observed no such pattern. On different simulation runs, a variety of different loci were the first (and last) to undergo substitution.

In addition to the above, we have tested how well Eq. (6) operates by determining the mean time interval between substitutions, from the numerical simulations, and the prediction of the equation that this time interval is $\rho^{-1} = 2m/\alpha$. Results were calculated by averaging the nine time intervals between substitutions and are summarized in Table 1.

In order to see the stochastic effects of genetic drift and if there any significant consequences of different rates-of-change of the optimal phenotypic value, α we have *centred* all allele frequencies around the *time of substitution*, t_{subs} (i.e. the time where $p = 1/2$). These are illustrated in Fig. 8 and to produce the figure, all 10 centred frequency profiles were, for a single value of α, centred and also numerically averaged.

It is evident from Fig. 8 that there is a reasonable level of similarity of the centred frequency profiles.

6. Conclusion

In this work we have investigated how a changing optimal phenotypic value of a quantitative trait affects the process of substitution at the different loci controlling the trait. During the process of substitution of the two alleles at any locus, the genetic variance is greatly enhanced above levels associated with mutation selection balance, even taking into account the range of genetic variances that are possible (Barton, 1986).

A striking feature of the numerical simulation for a finite population, is how periodically the substitutions occur (Fig. 7). This is far more regular than if there is a little heterogeneity in the properties of loci, in an infinite population (Fig. 4). The regularity associated with a finite population leads to the maximum value of the genetic variance (which fluctuates over time) being lower than the maximum value in an infinite population, because with an infinite population, a number of substitutions occur very close to each other and their combined contributions to the

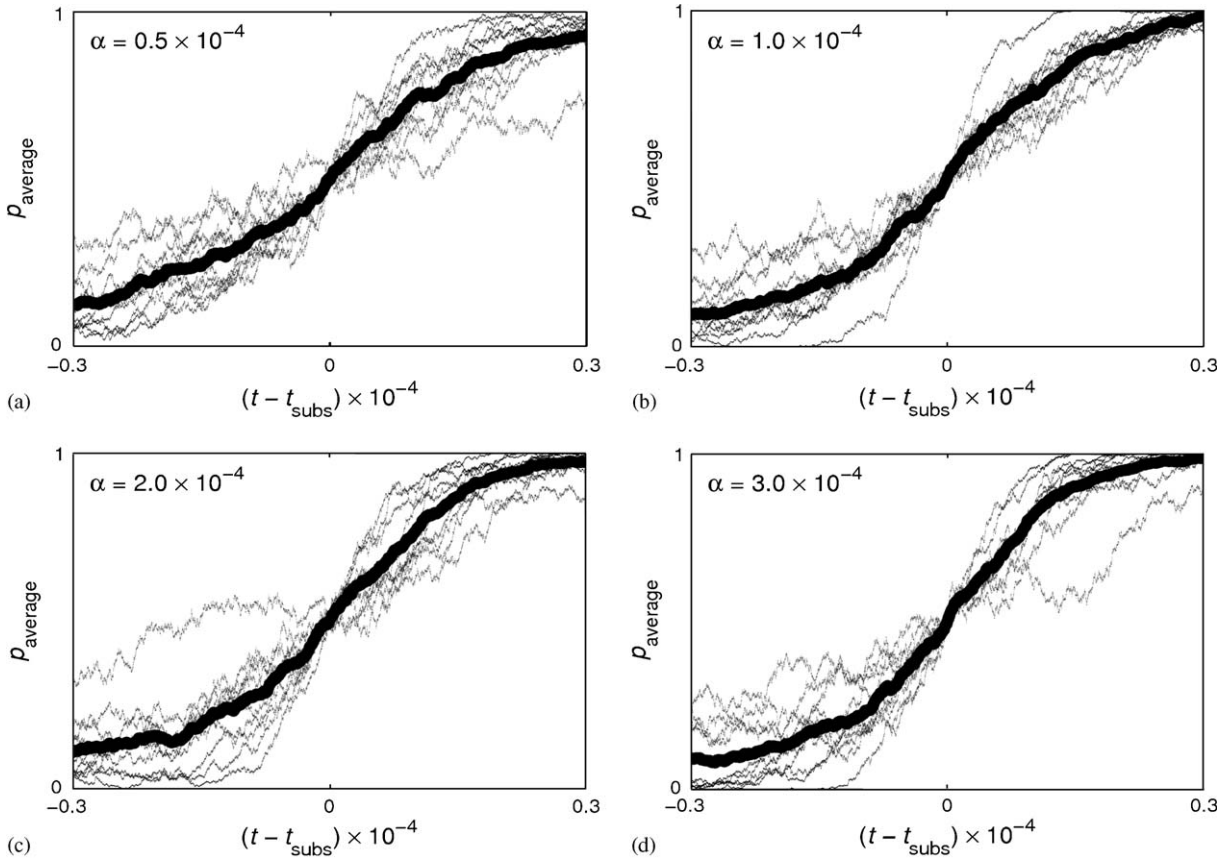

Fig. 8. To obtain the panels making up this figure, all allele frequencies (that were obtained from numerical simulation, as described in Fig. 5) were centred around the time, t_{subs}, when a substitution occurred. The centred allele frequencies are plotted as faint lines against the difference in time from the substitution time, t_{subs}. The thick black line, $p_{average}$, is the average of the centred allele frequencies and represents the frequency history associated with a typical substitution.

genetic variance add, under the approximation of linkage equilibrium. Indeed if substitutions are well separated in time, the maximum genetic variance associated with a substitution, at just one locus, is $m^2/2$, where m is the scale of allelic effects associated with the locus. Because mutation, when it occurs, results in an allele of opposite effect, compared with the pre-mutated allele, the quantity m^2 can also be interpreted as the variance in mutant effects. The value $m \simeq 0.2$ is often taken as a "typical" value and used in published work; it follows from Lande's (1983) extrapolation of the data of Russell et al. (1963). Accordingly, a maximum genetic variance associated with a single, non-overlapping, substitution is ~0.02. If substitutions are not well separated in time, or if the duration of a substitution is comparable with the time interval between substitutions, the genetic variance can be greater than $m^2/2$. In Fig. 7, the genetic variance fluctuates (over periods of thousands of generations) but is in the vicinity of 0.05 for much of the time and given the reasonable separation of substitutions, this value of the genetic variance arises from the appreciable duration of the substitutions; see Fig. 8.

In a previous work a related model was analytically investigated for a very large population with a continuum of alleles possible at every locus (Waxman and Peck, 1999). In Eq. (A13) of that work an approximation was given for the genetic variance, which, in the notation of the present work reads

$$V_G \simeq \sqrt{2n\alpha m V_s}\left[8\ln\left(\frac{m^2}{2uV_s}\right)\right]^{-1/4} \tag{7}$$

(we do not distinguish here between V_s and $V_s - 1$). It is interesting to compare this value, with the value that has been observed in the present work. Using the parameter values $n = 10$, $\alpha = 10^{-4}$, $V_s = 21$, $u = 10^{-5}$ we find Eq. (7) yields a value of $V_G \simeq 0.037$, which is in the vicinity of genetic variances we have obtained here from numerical simulation, from populations with 10 000 individuals.

We finish, by returning to the recent work we mentioned in the Introduction, by Gardner et al. (2004), in which they investigated the genetic variation for total fitness by carrying out experiments on replicate, caged populations of *Drosophila melanogaster*. They found results for variation in relative fitnesses over time that "could be due to subtle changes in external environment common to all cages." Indeed, they also stated that the "high variability we see is incompatible with the "classical" view, in which genetic variation is maintained by an equilibrium between deleterious mutations and selection". As we have already

stated in the Introduction of the present paper, knowledge of the strength of selection and the size of mutation rates may not, alone, be sufficient to predict the level of genetic variance of a population. Thus the results we have presented here may have some bearing on the findings of Gardner et al. (2004), since, for example, environmental change may have a significant influence. We note that since environmental change, in the form of a moving fitness optimum induces substitutions (i.e. influences the rate of evolution) which in turn causes enhancements in the genetic variance, it would be interesting to see if a value of the rate of change of the optimal phenotypic value, α, exists that is simultaneously compatible with observed levels of genetic variance *and* observed rates of evolution (the latter, via Eq. (6)).

Acknowledgements

We dedicate this work to the memory of John Maynard Smith (JMS). This work, like so many of the problems of current interest in evolutionary biology can be traced back to interests of JMS. We very much regret that he is not here to share his time and his insights with us.

It is a pleasure to thank Adam Eyre-Walker, Joel Peck and John Welch, who were all colleagues of JMS, and Yoshi Sato, for helpful discussions which have greatly improved this work. We also thank two anonymous referees, for their helpful comments. One of us (YB) was supported by an Alβan Scholarship of the European Union, the other (DW) was supported by the Leverhulme Trust.

Appendix A

In this Appendix, we determine when the quantity $\ln(E[w(G)])$ can be accurately approximated by $E[\ln(w(G))]$.

We proceed by first determining $E[w(G)]$. Using an identity following from a Gaussian integral, we can write the fitness function of Eq. (2) as

$$w(G) = \sqrt{\frac{V_s}{2\pi}} \int_{-\infty}^{\infty} \exp[\mathrm{i}k(G - Z_{opt}) - V_s k^2/2]\,\mathrm{d}k. \tag{8}$$

To find $E[w(G)]$ from this result, it is necessary to evaluate $E[\exp(\mathrm{i}kG)]$, which can be found in closed form, assuming Hardy Weinberg and Linkage equilibrium. Using Eq. (1) we have

$$\begin{aligned} E[\exp(\mathrm{i}kG)] &= \prod_{j=1}^{n} E[\exp(\mathrm{i}km_j(x_j + y_j)/2] \\ &= \prod_{j=1}^{n} \left[\left(p_j \mathrm{e}^{ikm_j/2} + q_j \mathrm{e}^{-ikm_j/2} \right)^2 \right] \\ &= \exp\left(2 \sum_{j=1}^{n} \ln(p_j \mathrm{e}^{\mathrm{i}km_j/2} + q_j \mathrm{e}^{-\mathrm{i}km_j/2}) \right). \end{aligned} \tag{9}$$

We note that the factor $\exp(-V_s k^2/2)$ in Eq. (8) only allows values of k satisfying $|k| \lesssim 1/\sqrt{V_s}$ to contribute to the integral. Accordingly, as long as $m_j/\sqrt{V_s} \ll 1$, for all j, we can expand the exponentials in Eq. (9) and keeping terms in the exponent to quadratic order in k yields

$$E[\exp(\mathrm{i}kG)] \simeq \exp(\mathrm{i}kE[G] - k^2 V_G/2), \tag{10}$$

where $E[G]$ and V_G are the mean and variance of G: $E[G] = 2\sum_{j=1}^{n} m_j(p_j - \frac{1}{2})$, $V_G = 2\sum_{j=1}^{n} m_j^2 p_j(1 - p_j)$. Using Eq. (10) in Eq. (8) quickly yields $\ln(E[w(G)]) \simeq -(E[G] - Z_{opt})^2/[2(V_s + V_G)] - \frac{1}{2}\ln(1 + V_G/V_s)$. Thus providing $V_G/V_s \ll 1$ and if, for all t of interest, $(E[G] - Z_{opt})^2 \ll V_s$, we have

$$\begin{aligned} \ln(E[w(G)]) &\simeq -\frac{(E[G] - Z_{opt})^2 + V_G}{2V_s} \\ &\equiv E[\ln(w(G))]. \end{aligned} \tag{11}$$

We have thus established when $\ln(E[w(G)])$ may be approximated by the analytically more convenient quantity $E[\ln(w(G))]$.

References

Barton, N.H., 1986. The maintenance of polygenic variation through a balance between mutation and stabilising selection. Genet. Res. (Camb.) 47, 209–216.

Barton, N.H., 1992. On the spread of new gene combinations in the third phase of Wright's shifting balance. Evolution 46, 551–557.

Barton, N.H., Shpak, M., 2000. The stability of symmetrical solutions to polygenic models. Theor. Popul. Biol. 57, 249–264.

Broom, M., Tang, Q., Waxman, D., 2003. Mathematical analysis of a model describing evolution of an asexual population in a changing environment. Math. Biosci. 186, 93–108.

Bulmer, M.G., 1974. Linkage disequilibrium and genetic variability. Genet. Res. (Camb.) 23, 281–289.

Bulmer, M.G., 1989. Maintenance of genetic variability by mutation-selection balance: a child's guide through the jungle. Genome 31, 761–767.

Bürger, R., 1999. Evolution of genetic variability and the advantage of sex and recombination in changing environments. Genetics 153, 1055–1069.

Bürger, R., Lynch, M., 1995. Evolution and extinction in a changing environment: a quantitative-genetic analysis. Evolution 49, 151–163.

Charlesworth, B., 1993. Directional selection and the evolution of sex and recombination. Genet. Res. (Camb.) 61, 205–224.

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

Gardner, M.P., Fowler, K., Barton, N.H., Partridge, L., 2004. Genetic variation for total fitness in Drosophila melanogaster: complex yet replicable patterns. Genetics 169, 1553–1571.

Gavrilets, S., 1997. Coevolutionary chase in exploiter-victim systems with polygenic characters. J. Theor. Biol. 186, 527–534.

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

Lande, R., 1983. The response to selection on major and minor mutations affecting a metrical trait. Heredity 50, 47–65.

Lynch, M., Walsh, B., 1998. Genetics and Analysis of Quantitative Traits. Sinauer, Sunderland, MA.

Maynard Smith, J., 1978. The Evolution of Sex. Cambridge University Press, Cambridge, UK.

Maynard Smith, J., 1988. Selection for recombination in a polygenic model—the mechanism. Genet. Res. (Camb.) 51, 59–63.

Russell, W.A., Sprague, G.F., Penny, H.L., 1963. Mutations affecting quantitative characters in long-time inbred lines of maize. Crop Sci. 3, 175–178.

Shpak, M., Kondrashov, A.S., 1999. Applicability of the hypergeometric phenotypic model to haploid and diploid populations. Evolution 53, 600–604.

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

Turelli, M., Barton, N.H., 1990. Dynamics of polygenic characters under selection. Theor. Popul. Biol. 38, 1–57.

Van Valen, L., 1973. A new evolutionary law. Evolutionary Theory 1, 1–30.

Waxman, D., 2000. Dynamics of a multilocus trait. Select. Mol. Genes, Memes 1, 229–245.

Waxman, D., Peck, J.R., 1999. Sex and adaptation in a changing environment. Genetics 153, 1041–1053.

Waxman, D., Peck, J.R., 2003. The anomalous effects of biased mutation. Genetics 164, 1615–1626.

Waxman, D., Welch, J.J., 2003. Non-equivalent loci and mutation-selection balance. Theor. Popul. Biol. 63, 339–345.

Welch, J.J., Waxman, D., 2002. Non-equivalent loci and the distribution of mutant effects. Genetics 161, 897–904.

Wright, S., 1935. The analysis of variance and the correlation between relatives with respect to deviations from an optimum. J. Genet. 30, 243–256.