

J.R. Miller · M.C. Pugh · M.B. Hamilton

## **A finite locus effect diffusion model for the evolution of a quantitative trait**

Received: 20 September 2005 / Revised version: 2 December 2005 /  
Published online: 7 February 2006 – © Springer-Verlag 2006

**Abstract.** A diffusion model is constructed for the joint distribution of absolute locus effect sizes and allele frequencies for loci contributing to an additive quantitative trait under selection in a haploid, panmictic population. The model is designed to approximate a discrete model exactly in the limit as both population size and the number of loci affecting the trait tend to infinity. For the case when all loci have the same absolute effect size, formal multiple-timescale asymptotics are used to predict the long-time response of the population trait mean to selection. For the case where loci can take on either of two distinct effect sizes, not necessarily with equal probability, numerical solutions of the system indicate that response to selection of a quantitative trait is insensitive to the variability of the distribution of effect sizes when mutation is negligible.

---

### **1. Introduction**

A quantitative trait is a continuous random variable. Examples include the height of a human, the oil content of a corn plant, and the number of bristles on the abdomen of a fruit fly. The value of a quantitative trait in an individual is generally determined by contributions from numerous loci (genes), as well as environmental factors and genetic environmental interactions. As a result, the evolutionary dynamics of quantitative traits are complicated to treat mathematically, and numerous models of such traits have been developed and analyzed.

One important class of models focuses on individual loci (Quantitative Trait Loci, or QTL) contributing to a given trait by directly analyzing haplotype or genotype frequencies. These include both two-locus ([35] [4] [15] [16] [5]) and multilocus models ([19] [18] [1] [14] [17]). Many of these studies aim to determine conditions for maintenance or non-maintenance of genetic variability at equilibrium in a large population under stabilizing selection [16] [35] [1] [19] [5] [17]; the dynamics of the trait mean and variance are considered in a smaller number of studies, e.g. [16] [18] [15].

---

J.R. Miller: Department of Mathematics, Georgetown University, Washington, DC 20057-1233, USA. e-mail: jrm32@georgetown.edu

M.C. Pugh: Department of Mathematics, University of Toronto, Toronto, ON M5S 3G3, Canada

M.B. Hamilton: Department of Biology, Georgetown University, Washington, DC 20057-1229, USA

A host of other models have been used to analyze the moments of the distribution of phenotypes and breeding values in a population, under a wide variety of assumptions. The oldest such model is the infinitesimal model, the key assumption of which is that a large, effectively infinite number of equivalent loci each make a small, effectively infinitesimal contribution to the trait under study [12] [3]. Other models incorporating varying amounts of detail at the locus level may be kept tractable by assuming distributions of phenotypes, breeding values and/or allelic effects to be Gaussian [26] [25] [28], by assuming selection to be weak relative to other forces (e.g. recombination) [2] [32] [28] [35], or conversely by assuming selection to be strong [33] [17]. As with the two-locus and multilocus models mentioned above, much analysis of these models aims to predict response to selection or the amount of genetic variability maintained at equilibrium.

The long-term response to selection of a quantitative trait depends on numerous genetic details, including the initial allele frequencies at loci contributing to the trait, allelic effects (i.e. the size of the contributions of QTL to the trait), linkage between loci, and gene interactions [11] [34]. As discussed above, much theoretical work on long-term selection assumes that the population trait mean is at or near equilibrium and concerns the maintenance of genetic variation under mutation-selection balance (e.g. [26] [31]). A smaller body of work concerns the approach to a selection limit as existing genetic variation is depleted (e.g. [18], [16] [34] [22] [30] [7]). Here we address this latter problem; however, the model developed here should be modifiable to include mutation.

We present a diffusion model for the evolution of a quantitative trait. It is a “mesoscale” model in that it explicitly incorporates finite (i.e. non-infinitesimal) allelic effects, but does not allow the tracking of allele frequencies at individual QTL during evolution. The model leads to a partial differential equation (PDE) for the joint distribution of allele frequency and effect size; there is no assumption that the distributions of these two variables are independent, in contrast to the model in [22]. The strength of selection on individual loci depends on the fitness function chosen as well as on allele frequencies and can in general vary with time. When selection is absent, the model reduces to the classical Wright-Fisher diffusion equation of population genetics [8] [10].

The PDE in our model includes a nonlocal term which represents the trait mean at any given time as the integral of contributions from all loci. Since this integral evolves along with the distribution of allele frequencies and effect sizes, no assumption is made that the mean is at or near an optimum value. On the other hand, the coupling of loci through the trait mean limits the interpretation of the model. In (e.g.) the Wright-Fisher equation, we may interpret the solution either as a distribution of allele frequencies for one locus in multiple independent populations or as a distribution over multiple loci in a single population. In the present model only the latter interpretation is reasonable.

By treating the PDE as a perturbation of the Wright-Fisher model, it is possible to obtain asymptotic approximations for solutions and thereby to predict the dynamics of the trait mean from initial data. We carry out a formal analysis of this type for the case where all loci have equal effects. In particular, we obtain analytical expressions for the leading-order behavior of the trait mean, which include a

prediction of the selection limit as genetic variance is exhausted. The predictions are likely to be most accurate for a trait which is experiencing weak selection following a lengthy period of evolution in neutrality.

Below, we first sketch a derivation of the diffusion model as a limit of an underlying discrete stochastic process and discuss its interpretation in the context of that process. We develop multiple-time-scale asymptotic expressions to describe the long-time behavior of certain aspects of the model under weak selection, and check the validity of the asymptotics by comparison with numerical solutions of the full diffusion system. We then use numerical solutions to examine the influence of locus effect size variance on the dynamics of the trait mean, for specific models in which loci can take on either of two effect sizes.

## 2. The model

We now describe the discrete evolutionary model that our model, described below, is designed to approximate. The discrete model is of a single panmictic population of constant size  $N$ , with  $n$  unlinked, diallelic haploid loci contributing to a particular trait. We ignore any environmental contributions to the trait. We also assume strict additivity of allelic contributions to the trait value of individuals, i.e. that epistasis is absent.

The alleles at locus  $i$  are denoted by  $A_i$  and  $a_i$ , with the convention that  $A_i$  contributes to a higher trait value than  $a_i$ . The fraction of  $A_i$  alleles at the  $i$ th locus is denoted by  $p_i$ . The average effect of locus  $i$ , that is, one-half the mean difference in phenotype between haploid individuals carrying  $A_i$  and those carrying  $a_i$ , is denoted by  $\alpha_i$ , with units identical to those of the trait. Then under strict additivity the population phenotypic mean  $\bar{z}$  is given (up to an additive constant called the midpoint which we decree to be 0) by

$$\bar{z} = \sum_{i=1}^n p_i \alpha_i + (1 - p_i)(-\alpha_i) = \sum_{i=1}^n (2p_i - 1)\alpha_i. \quad (1)$$

All pairs of loci are in gametic phase equilibrium if the distributions of alleles at loci  $i$  and  $j$  are independent for  $i \neq j$ . If this holds and  $\bar{z}_{A_i}$  is the mean phenotype of individuals carrying the  $A_i$  allele (i.e. the genotypic value of the  $A_i$  allele) and  $\bar{z}_{a_i}$  is the mean phenotype of those carrying the  $a_i$  allele then  $2\alpha_i = \bar{z}_{A_i} - \bar{z}_{a_i}$ .

To model natural selection, we require a relative fitness function  $w(z)$ , giving the expected number of offspring for an individual with trait value  $z$  as a fraction of the expected number of offspring for an individual with the optimal trait value, denoted  $z_{opt}$ . The relative fitness function  $w(z)$  is assumed to satisfy the following conditions:

$$\begin{aligned} w(z) &\text{ depends only on } |z - z_{opt}| \\ w(z) &\geq 0 \quad \text{for all real } z \\ w(z_{opt}) &= 1. \end{aligned}$$

In what follows, we will employ a Gaussian fitness function

$$w(z) = e^{-(z-z_{opt})^2/2V_s}. \quad (2)$$

The coefficient  $1/2V_s$  in (2) determines the strength of selection; note that as  $V_s \rightarrow \infty$ , all phenotypes become equally fit.

We require an expression for the probability  $p'_i$  that an individual in generation  $\tau + 1$  will carry allele  $A_i$ , given allele frequencies at all loci in generation  $\tau$ . This probability depends on the allele frequencies  $p_j(\tau)$ , and the effect sizes  $\alpha_j$  at all loci ( $j = 1, \dots, n$ ), as well as on the fitness function  $w$ . To derive an expression for  $p'_i$ , we start by noting that the population contains a total of  $Np_i(\tau)$  alleles of type  $A_i$  and  $N(1 - p_i(\tau))$  alleles of type  $a_i$  after  $\tau$  generations. A sample, with replacement, of size  $N$  from the existing alleles will be passed to the next generation ( $\tau + 1$ ). Let  $w_{A_i}$  and  $\bar{z}_{A_i}$  (respectively,  $w_{a_i}$  and  $\bar{z}_{a_i}$ ) denote the mean fitness and the mean phenotype of individuals with allele  $A_i$  ( $a_i$ ). Then the expected proportion  $p'_i$  of  $A_i$  alleles in generation  $\tau + 1$  must be proportional to both  $Np_i$  and  $w_{A_i}$ ; similar considerations apply to  $q'_i$ , the expected proportion of  $a_i$  alleles. Since the total number of alleles must equal  $N$  in every generation, it follows that

$$p'_i = \frac{p_i w_{A_i}}{p_i w_{A_i} + (1 - p_i) w_{a_i}} \quad \text{and} \quad q'_i = 1 - p'_i. \quad (3)$$

To compute  $p'_i$ , one needs to evaluate  $w_{A_i}$  and  $w_{a_i}$ . To evaluate  $w_{A_i}$ , say, one needs to identify each member of the population with the  $A_i$  allele and compute that member's fitness. The average of those fitness values is  $w_{A_i}$ . This calculation would have to be done for each locus and at each generation. To simplify computations, we do not use the exact expression (3). Rather, we make the assumption that all pairs of loci are in gametic phase equilibrium. This allows us to use the approximate expression

$$p'_i \approx p_i{}^{approx} = \frac{p_i w(\bar{z}_{A_i})}{p_i w(\bar{z}_{A_i}) + (1 - p_i) w(\bar{z}_{a_i})}, \quad (4)$$

valid in the limit of weak selection. A justification of this approximation is given in Appendix B. Using equation (1) for the overall trait mean  $\bar{z}$ , it also follows under gametic phase equilibrium that

$$\begin{aligned} \bar{z}_{A_i} &= \bar{z} - (2p_i - 1)\alpha_i + \alpha_i = \bar{z} + 2(1 - p_i)\alpha_i, \\ \bar{z}_{a_i} &= \bar{z} - (2p_i - 1)\alpha_i - \alpha_i = \bar{z} - 2p_i\alpha_i. \end{aligned} \quad (5)$$

Substituting (5) into (4) yields

$$p'_i = \frac{p_i w(\bar{z} + 2(1 - p_i)\alpha_i)}{p_i w(\bar{z} + 2(1 - p_i)\alpha_i) + (1 - p_i) w(\bar{z} - 2p_i\alpha_i)}. \quad (6)$$

This expression is easier to compute than (3) since it depends only on the allele frequencies and on the effect sizes.

Evolution in the discrete model outlined above is a stochastic process, in which the number of  $A_i$  alleles in generation  $\tau + 1$  is a binomial random variable with  $N$  trials and probability  $p'_i$  of success (i.e. of drawing an  $A_i$  allele) at each trial. We wish to derive a deterministic model for the expected or mean behavior of this process. It should be possible, using standard techniques, to derive a (continuous-time) system of  $n$  ordinary differential equations (ODEs) that approximates the expected

behavior of each allele frequency in the discrete model [10]. However, such a system will quickly become unwieldy as  $n$  increases. Instead, we now describe a diffusion model that explicitly accounts for QTL effects and allele frequencies but does not track allele frequencies at the  $n$  individual loci.

The key idea in the derivation of the diffusion model is that the trait mean  $\bar{z}$  can be expressed as a sum not over loci as in (1), but rather over effect sizes and allele frequencies. To do so, let  $\Phi(p, \alpha, \tau)$  denote the number of loci with allele frequency  $p$  and effect size  $\alpha$  after  $\tau$  generations. Then we have

$$\bar{z}(\tau) = \sum_p \sum_\alpha \alpha(2p - 1)\Phi(p, \alpha, \tau). \tag{7}$$

The summation in (7) amounts to an indexing of loci by allele frequency and effect size, rather than by arbitrary integers.

A standard formal derivation (as in, e.g., [10]) of a nonlocal diffusive PDE from the discrete stochastic process described above is given in Appendix A. This model is obtained in terms of dimensionless variables

$$\tilde{V}_s = \frac{V_s n}{N z_0^2}, \quad \tilde{\alpha} = \frac{\alpha n}{z_0}, \quad \tilde{z}_{opt} = \frac{z_{opt}}{z_0}, \quad \tilde{z} = \frac{\bar{z}}{z_0}, \quad t = \frac{\tau}{N}, \tag{8}$$

where  $z_0$  is a characteristic trait scale (e.g. the phenotypic standard deviation of the founding population). It is expected to be valid in the limit as both the population size  $N$  and the number of loci  $n$  become large, with  $n/N$  being assumed small and with the selection coefficient  $1/2V_s$  being on the order of  $n/N$ ; thus selection is assumed to be weak. (Note that in nearly all biologically relevant situations  $N \gg n$ .)

In the following, we will drop tildes and not explicitly distinguish between the dimensional variables and the dimensionless variables. We will refer to  $\bar{z}$  (henceforth  $\bar{z}$ ) as the trait mean, for example.

The dependent variables in the model are functions  $\phi(p, \alpha, t)$  (defined for  $p \in (0, 1)$ ),  $\phi_0(\alpha, t)$  and  $\phi_1(\alpha, t)$ . Together,  $\phi(p, \alpha, t)$ ,  $\phi_0(\alpha, t)\delta_0(p)$  and  $\phi_1(\alpha, t)\delta_1(p)$  form a joint probability density for allele frequency  $p \in [0, 1]$  and effect size  $\alpha$  at time  $t$ ; the point masses represent the proportions of loci that have become fixed at  $p = 0$  and  $p = 1$ . The trait mean  $\bar{z}$  is defined by

$$z(t) = \sum_\alpha \int_0^1 (2p - 1) \alpha \phi(p, \alpha, t) dp + \sum_\alpha \alpha \phi_1(\alpha, t) - \alpha \phi_0(\alpha, t). \tag{9}$$

The dependent variable  $\phi$  satisfies the PDE

$$\phi_t = -(M\phi)_p + \frac{1}{2}(V\phi)_{pp}, \tag{10}$$

where the coefficients  $M(p, \alpha, \bar{z}(t))$  and  $V(p, \alpha, \bar{z}(t))$  are determined by the selection coefficient  $1/2V_s$  and  $\bar{z}(t)$ , the trait mean at time  $t$ :

$$M = \frac{2}{V_s} \alpha p(1 - p)(z_{opt} - \bar{z}(t)), \tag{11}$$

$$V = p(1 - p). \tag{12}$$

Since  $\bar{z}$  is defined as an integral involving  $\phi$ , the PDE (10) is nonlocal. Finally, the point masses  $\phi_0$  and  $\phi_1$  are given by

$$\phi_0(\alpha, t) = \phi_0(\alpha, 0) - \int_0^t \sum_{\alpha} \lim_{p \rightarrow 0^+} \left( [M\phi(p, \alpha, s) - \frac{1}{2} \partial_p (V\phi(p, \alpha, s))] \right) ds \quad (13)$$

$$\phi_1(\alpha, t) = \phi_1(\alpha, 0) + \int_0^t \sum_{\alpha} \lim_{p \rightarrow 1^-} \left( [M\phi(p, \alpha, s) - \frac{1}{2} \partial_p (V\phi(p, \alpha, s))] \right) ds. \quad (14)$$

A few remarks on the model (9)–(13) are in order. First, the PDE (10) involves no derivatives with respect to  $\alpha$ . This is because in our model a locus will not change its effect size as time passes, but allele frequencies at a locus can and in general will change. For example, suppose that the  $A$  allele frequency at a locus with effect size  $\alpha$  increases by  $\Delta p$  in time  $\Delta t$ ; this would increase the number of loci with allele frequency  $p + \Delta p$  and effect size  $\alpha$  at time  $t + \Delta t$ , but would not affect the distribution of allele frequencies for loci with effect size different from  $\alpha$ . The coupling between loci with *different* effect sizes arises only through the trait mean  $\bar{z}(t)$ , which (as equation (9) indicates) is a type of average over all loci.

Second, if a QTL has become fixed (i.e. has allele frequency  $p = 0$  or  $1$ ) at a particular time then it will remain fixed — there is no mutation in the present model. This may be thought of as an absorbing boundary condition for the PDE (10). Since loci that have become fixed continue to contribute to the trait mean, it is necessary to keep track of the “mass” that has been absorbed through the boundaries  $p = 0$  and  $p = 1$ . This motivates the definition (13) of the point masses  $\phi_0(\alpha, t)$  and  $\phi_1(\alpha, t)$ : i.e., the proportion of loci that have become fixed at  $p = 0$  or  $p = 1$  between times  $0$  and  $t$  is given by integrating the flux through the boundary over time.

The trait mean is a population-level quantity. And so, because of the coupling of loci through the trait mean in the coefficients (11), we view the diffusion (10) as describing the behavior of a single population. In particular, the conditional probability integral

$$\frac{\int_{p=a}^b \phi(p, \alpha, t) dp}{\int_{p=0}^1 \phi(p, \alpha, t) dp}$$

should be interpreted as describing the probability that a locus with effect size  $\alpha$  in *this single population* will have an allele frequency in  $(a, b)$  at time  $t$ . This interpretation is quite different from that in the classical models, where  $\phi$  can be viewed as a probability distribution over replicate populations. Moreover, it means that the effects of genetic sampling (i.e. of sampling due to finite population size) and of sampling due to the fact that only finitely many loci contribute to the trait are not included in the present model. This point is discussed further in section 6.

In addition to the trait mean  $\bar{z}(t)$ , another interesting time-dependent quantity is

$$S^2(t) = \sum_{\alpha} \int_0^1 4\alpha^2 p(1-p)\phi(p, \alpha, t) dp. \quad (15)$$

This is the nondimensionalized total additive genetic variance,  $S^2 = n\sigma_G^2/z_0^2$ , where  $\sigma_G^2$  is the (dimensional) total additive genetic variance and  $z_0$  and  $n$  are the characteristic trait scale and the number of loci contributing to the trait. Because the model does not include mutation, the absorbing boundary conditions mean that we expect allele frequencies at all loci to eventually go to 0 or 1:  $\phi(p, \alpha, t) \rightarrow 0$  as  $t \rightarrow \infty$ . Therefore, in the long-time limit the trait mean  $\bar{z}(t)$  should tend to a constant value and the total additive genetic variance  $S^2(t)$  should tend to zero.

### 3. Numerical simulations

In this section, we present numerical simulations of solutions of the system (9)–(13). The absorbing boundary conditions pose an immediate computational difficulty. Equation (10) shares with classical models the property that in some regimes solutions are expected to grow rapidly near the absorbing boundaries. For this reason, we compute solutions of a closely related initial value problem and then transform these solutions to solutions of (9)–(13).

We define the new variable

$$\psi(p, \alpha, t) := V(p, \alpha, \bar{z}(t))\phi(p, \alpha, t)$$

for  $p \in (0, 1)$ . Equation (10) then becomes

$$\begin{aligned} \psi_t(p, \alpha, t) = & \frac{\psi(p, \alpha, t)}{V(p, \alpha, \bar{z}(t))} \frac{d}{dt} V(p, \alpha, \bar{z}(t)) \\ & - V(p, \alpha, \bar{z}(t)) \left( \frac{M(p, \alpha, \bar{z}(t))}{V(p, \alpha, \bar{z}(t))} \psi(p, \alpha, t) \right)_p \\ & + \frac{1}{2} V(p, \alpha, \bar{z}(t)) \psi_{pp}(p, \alpha, t). \end{aligned} \tag{16}$$

The numerical simulations solve the system (16), where the trait mean  $\bar{z}(t)$  is computed with  $\psi/V$  wherever  $\phi$  appears in the formula (9) with  $p \in (0, 1)$ . Given initial data  $\phi(p, \alpha, 0)$ , we compute  $\bar{z}(0)$  and then  $V(p, \alpha, 0)$  to determine the initial data  $\psi(p, \alpha, 0)$ . We impose Dirichlet boundary conditions on the solution:  $\psi(0, \alpha, t) = \psi(1, \alpha, t) = 0$  for all  $\alpha$  and  $t$ . This is reasonable since  $\phi$  is bounded above by a probability distribution, and so we should have  $\int_0^1 \phi dp < \infty$  for all  $t$ . Since  $M$  and  $V$  vanish linearly at the boundary, we find that  $\phi$  should satisfy

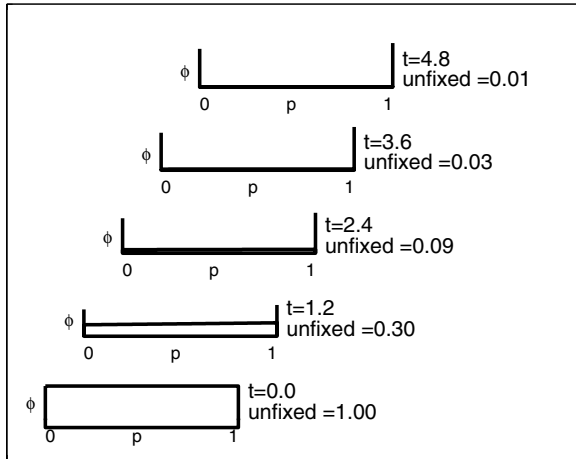
$$\lim_{p \rightarrow 0^+} M\phi = \lim_{p \rightarrow 1^-} M\phi = \lim_{p \rightarrow 0^+} V\phi = \lim_{p \rightarrow 1^-} V\phi = 0. \tag{17}$$

As an additional diagnostic, we compute the total additive genetic variance  $S^2(t)$  at each timestep.

The scheme is fourth-order accurate in space, using finite differences on a fixed, nonuniform mesh to approximate the  $p$ -derivatives. The mesh was chosen to have higher resolution near  $p = 0$  and  $p = 1$ . The time-stepping is second-order accurate in time. It uses a Crank-Nicolson time-stepping for the diffusive term, extrapolation for the source term, and an extrapolated leapfrog method for the advection term. This computes the solution at the interior meshpoints. To compute the density of

loci being fixed at  $p = 0$  and  $p = 1$  (the mass of the delta functions) a pair of ODEs were solved using Richardson extrapolation. The ODEs are solved simultaneously with the diffusion equation (16) because fixed loci continue to contribute to the trait mean  $\bar{z}(t)$ , which in turn appears in the coefficients of the diffusion equation. The code was written and executed in Matlab on a PC; the code is available upon request.

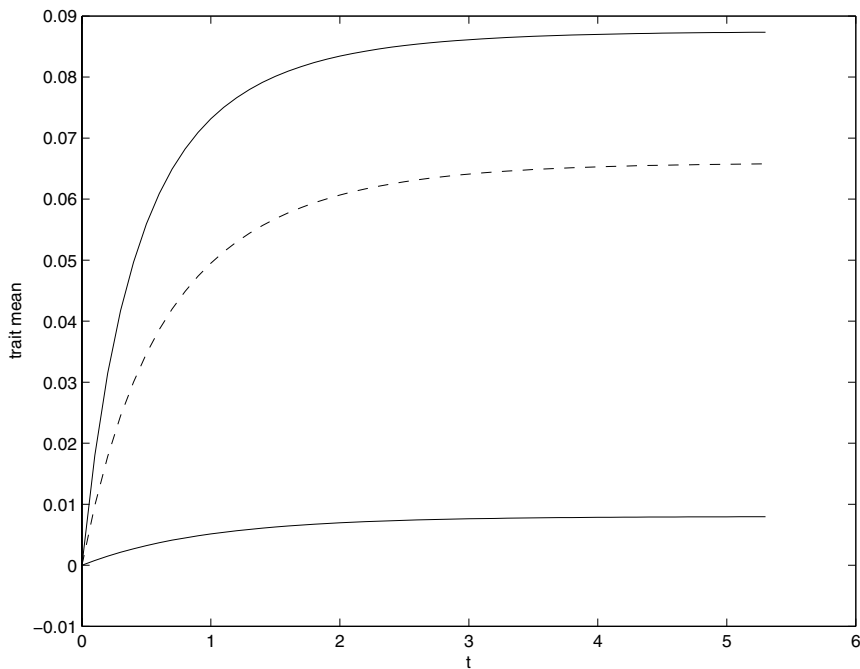
In Figure 1, we present the solution to (9)–(13) using a single value of the QTL effect size  $\alpha$  (specifically,  $\alpha = 1.25$ ), with uniform unfixed initial data ( $\phi(x, \alpha, 0) = 1$  for  $0 < p < 1$ ) and no initial fixation ( $\phi_0(\alpha, 0) = \phi_1(\alpha, 0) = 0$ ). The free parameters are taken as  $V_s = 2$  and  $z_{opt} = .1$ . The simulation was continued until the time at which just 0.5% of loci remained unfixed:  $t_f = 5.3$ , corresponding to  $5.3N$  generations. The solution  $\phi$  is shown at five approximately equally spaced times between 0 and  $t_f$ . Note that the distribution of allele frequencies  $\phi(x)$  is slightly skewed toward  $p = 1$  for all times  $t > 0$  computed. This skewness reflects the action of selection. The initial data is symmetrical about  $p = 1/2$  and so the initial trait mean is zero. Therefore, if the optimum trait value  $z_{opt}$  were equal to 0



**Fig. 1.** The solution to the system (9)–(13) shown at five times. We assumed only one value of the QTL effect size:  $\alpha = 1.25$ , and chose parameters  $1/2V_s = 1$ ,  $z_{opt} = .1$ . The initial data is  $\phi(p, 1.25, 0) = 1$  for  $0 < p < 1$ , and  $\phi_0(1.25, 0) = \phi_1(1.25, 0) = 0$ . For five times, we present the solution  $\phi(p, 1.25, t)$  (plotted as a graph) and the fixations  $\phi_0(1.25, t)$  and  $\phi_1(1.25, t)$  (presented as spikes located at  $p = 0, 1$  of the respective heights). The front-most plot is at time  $t = 0$ ; 100% of the loci are unfixed. The plot one further back is at  $t = 1.2$ ; 30% of the loci are unfixed, 32% and 38% are fixed at  $p = 0$  and  $p = 1$  respectively. The third plot is at  $t = 2.4$ ; 9% of the loci are unfixed, 42% and 49% are fixed at  $p = 0$  and  $p = 1$  respectively. The fourth plot is at  $t = 3.6$ ; 2.7% of the loci are unfixed, 45% and 52% are fixed at  $p = 0$  and  $p = 1$  respectively. The fifth plot is at  $t = 4.8$ ; 0.8% of the loci are unfixed, 46% and 53% are fixed at  $p = 0$  and  $p = 1$  respectively.

there would be no selective pressure to change the trait mean, only to reduce trait variance. The solution would thus remain symmetrical about  $p = 1/2$  at all times. Since the optimum trait value  $z_{opt}$  is in fact positive, however, alleles that increase an individual's phenotypic value are favored over alleles that decrease it, resulting in skewness of the distribution  $\phi$  with a bias towards fixing at  $p = 1$ . If  $z_{opt}$  were negative, the bias would be towards  $p = 0$ . As will be seen in Section 4, if no selection were present the system would become linear and, since the horizontal initial data is precisely the leading eigenfunction of the relevant linear operator, no higher modes would be excited and thus no skewness would develop. The skewness of the solution is thus a nonlinear effect due to coupling of loci through the trait mean  $\bar{z}$  and the fitness function  $w$ . When the simulation was rerun with  $z_{opt} = 1$  and all other parameters and initial data the same as for Figure 1, more pronounced skewness as well as concavity could be seen, reflecting the stronger action of selection. In addition, it took only  $4.4N$  generations to reduce the percentage of unfixed loci to 0.5.

Figure 2 plots the trait mean  $\bar{z}(t)$  as a function of time for the  $\alpha = 1.25$  simulation (see Figure 1) and for an analogous simulation with  $\alpha = 0.25$ . As expected,  $\bar{z}(t)$

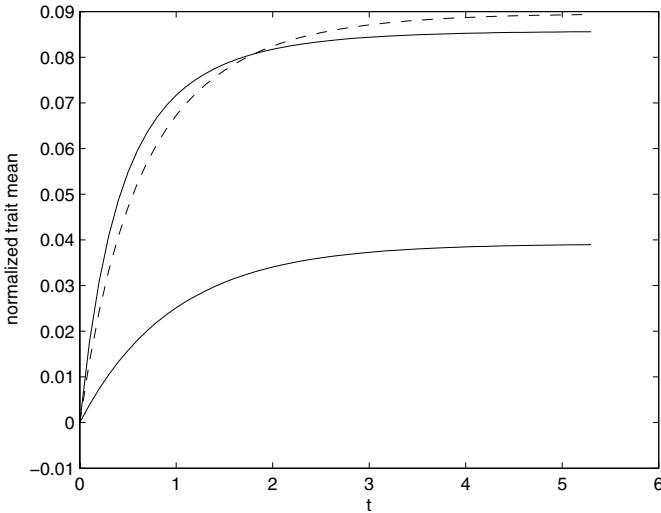


**Fig. 2.** The trait mean  $\bar{z}(t)$  is plotted as a function of time for three solutions. The upper solid line corresponds to the QTL effect size  $\alpha = 1.25$  simulation shown in Figure 1. The bottom solid line corresponds to the  $\alpha = 0.25$  simulation. The dashed line corresponds to a simulation which has two values of  $\alpha$ :  $\alpha = 1.25$  and  $\alpha = 0.25$ . The  $V_s$  and  $z_{opt}$  are as in the other two simulations and the initial data is  $\phi(p, 1.25, 0) = 1/2$  (and  $\phi_0(1.25, 0) = \phi_1(1.25, 0) = 0$ ) and  $\phi(p, 0.25, 0) = 1/2$  (and  $\phi_0(0.25, 0) = \phi_1(0.25, 0) = 0$ ).

approaches an equilibrium much closer to the optimum value  $z_{opt} = .1$  for the larger value of  $\alpha$ . This partly reflects the stronger action of selection for  $\alpha = 1.25$ , but also reflects the fact that the total additive genetic variance was greater at  $t = 0$  when  $\alpha = 1.25$  than when  $\alpha = 0.25$ :  $S^2(0) = 4(1.25)^2/6$  versus  $S^2(0) = 4(0.25)^2/6$ . Thus the solution with  $\alpha = 1.25$  advanced further before “running out of steam” at 99.5% fixation.

The dotted line in Figure 2 presents  $\bar{z}(t)$  for the solution to (9)–(13) with half of all loci having effect size  $\alpha = 1.25$  and half having  $\alpha = 0.25$ . The initial data is uniform in  $p$  for both values of  $\alpha$  and the parameters  $1/2V_s$  and  $z_{opt}$  are unchanged. Nonlinear effects are also apparent in Figure 2: the trait mean  $\bar{z}(t)$  for the two- $\alpha$  case is much closer to the trait mean for  $\alpha = 1.25$  than to an average of the trait means for  $\alpha = 1.25$  and  $\alpha = 0.25$ . This is further demonstrated in Figure 3. Here, the response of the trait mean to selection was normalized by dividing  $\bar{z}(t)$  by the square root of the total additive genetic variance of the founding population and then plotted. The responses for the  $\alpha = 1.25$  simulation and the 50:50 simulation were very similar, while the response was less for the  $\alpha = .25$  simulation. This must reflect a weaker selection intensity for the latter case [11].

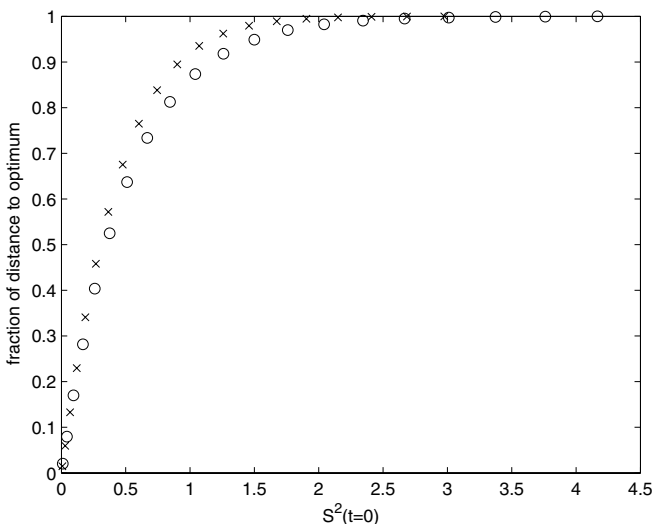
Since there is no mutation in the present model to replenish genetic variation, it is possible for the trait mean to fall short of the optimum as the initial supply of such variation is exhausted. Figure 4 illustrates how initial genetic variance and allele frequencies affect the advance of the trait mean toward its optimum; we note that in



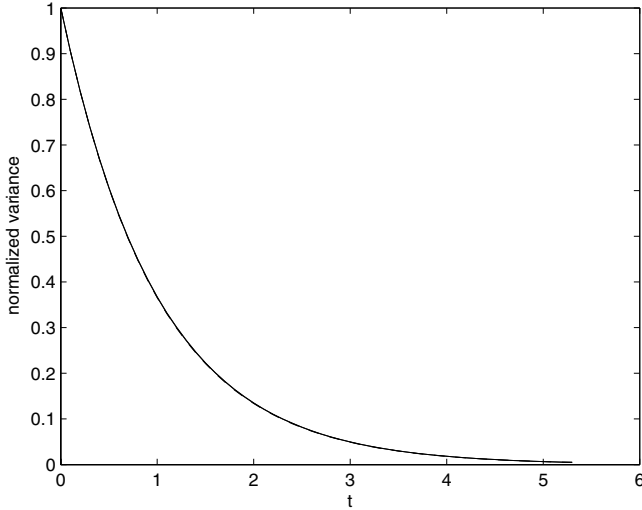
**Fig. 3.** The trait mean  $\bar{z}(t)$ , divided by the square root of  $S^2(t = 0)$ , is plotted as a function of time for three solutions. The upper solid line corresponds to the QTL effect size  $\alpha = 1.25$  simulation shown in Figure 1. The bottom solid line corresponds to the  $\alpha = 0.25$  simulation. The dashed line corresponds to a simulation which has two values of  $\alpha$ :  $\alpha = 1.25$  and  $\alpha = 0.25$ . Parameters, distributions of  $\alpha$  and initial data are as in Figure 2.

most of the numerical runs whose outcomes are depicted, initial genetic variance was insufficient to attain the optimum even in the absence of any drift. Here the fraction  $[\bar{z}(t_f) - z_{opt}] / [\bar{z}(0) - z_{opt}]$  of the initial distance to the optimum covered during the course of evolution (to 99.5% fixation) is plotted against initial genetic variance ( $S^2$ ) for runs where initial allele frequencies were distributed uniformly (data denoted by  $\circ$ ) or according to the distribution  $\phi(p, 0) = 5(1 - p)^4$  (data denoted by  $\times$ ). In Figure 4 we see that the curves for the two types of initial allele frequency distributions are close to each other and that there is a strong dependence of progress toward the optimum on initial variance when the latter is small. Thus for the conditions tested, initial genetic variance influenced the amount of progress made toward the optimum more than did initial allele frequency. This confirms results obtained using a different model by Hill and Rasbash [22].

Finally, Figure 5 plots the total additive genetic variance  $S^2$  (normalized by dividing by  $S^2(t = 0)$ ) as a function of  $t$  for the three numerical runs for which trait means are plotted in Figures 1 and 2. The three curves in Figure 5 are indistinguishable from one another and from the corresponding curve obtained from the same initial data with zero selection (replacing the Gaussian fitness function (2) with the constant function  $w(Z) \equiv 1$ ) suggesting that at the low selection strengths appropriate for the present model, variance dynamics are consistent with the exponential decay of heterozygosity expected at a locus under neutrality and random mating (this is a basic result in population genetics; see for example Chapter 7 of [20]).



**Fig. 4.** Progress toward the optimum as a fraction of initial distance from the optimum, plotted against initial genetic variance ( $S^2$ ). Runs with uniform initial data are denoted by  $\circ$ , runs with quartic initial data by  $\times$ . Values of  $\alpha$  ranged from .125 to 2.5 for each initial distribution.



**Fig. 5.** Normalized trait variance as a function of time for the solutions of the system (10)–(13) for which parameters, distributions of  $\alpha$  and initial data are as in 2 and 3. The three curves are indistinguishable from one another and from the corresponding curve obtained with zero selection.

At first these findings might appear to contradict those of Gavrillets and Hastings [18] (see also [6]), who presented simulations indicating that when  $N/n$  is large, the decay of genetic variance is not exponential (recall that  $N \gg n$  was assumed in our derivation of the PDE (10)). However, Gavrillets and Hastings assumed at least moderately strong selection and, more importantly, initial trait means close to the optimum. Referring to the derivation of the system (10)–(13) in Appendix A, we note that if the trait mean  $\bar{z}$  were equal to the optimum  $z_{opt}$  and higher-order terms were included in the derivation of the mean change in allele frequency  $M$  (A11), a different PDE would result. For this PDE it is possible to follow the procedure of Gavrillets and Hastings to derive variance dynamics like those predicted in [18] (details not shown).

#### 4. Multiple-scale expansions

Two biological processes are at work in the model: random drift and selection. In equation (10), random drift is modeled by the diffusive term  $(V\phi)_{pp}$  and selection is modeled by the advective term  $(M\phi)_p$ . Each process has its own characteristic timescale. The timescale of random drift may be taken to be 1, which serves as a diffusion coefficient in the definition of  $V$  (11). The timescale of selection is then determined by the value of  $\alpha/V_s$ , together with  $z_{opt}$ . To see this, suppose that  $z_{opt} = 0$ ; then  $\alpha/V_s$  functions as a speed of propagation in the advection term of equation (10). (We remind the reader that here  $V_s$  and  $\alpha$  are dimensionless variables.

Using the rescaling (8), the timescale of selection in terms of dimensional variables is  $\alpha N_{z_0}/V_s$ .)

If selection is weak relative to random drift,

$$\alpha/V_s \ll 1,$$

then there is a separation in timescales. This motivates the use of a two-timescale asymptotic expansion (see for example [23, §6]) to approximate the model dynamics in the case of weak selection. We now outline such an analysis for the case when all loci share a single effect size  $\alpha$ .

To begin, we use the slower timescale (selection) as an expansion parameter:

$$\epsilon = \alpha/V_s$$

and assume that there is a separation in timescales:  $\epsilon \ll 1$ . We introduce two time variables, one for each timescale:

$$\tau = t \quad T = \epsilon t$$

The distribution  $\phi$  and trait mean  $\bar{z}$  are then written in terms of the new time variables and expanded in  $\epsilon$ :

$$\begin{aligned} \phi(p, \tau, T) &= \phi_0(p, \tau, T) + \epsilon\phi_1(p, \tau, T) + \epsilon^2\phi_2(p, \tau, T) + H.O.T. \\ \bar{z}(\tau, T) &= \bar{z}_0(\tau, T) + \epsilon\bar{z}_1(\tau, T) + \epsilon^2\bar{z}_2(\tau, T) + H.O.T. \end{aligned} \tag{18}$$

“H.O.T.” is short-hand for “higher-order terms”. (Throughout this section, we are abusing notation in the following manner. In other sections,  $\phi_0(\alpha, t)$  and  $\phi_1(\alpha, t)$  represent fixation at  $p = 0, 1$ . In this section,  $\phi_0(p, \tau, T)$  and  $\phi_1(p, \tau, T)$  refer to functions that are part of the multiple scale expansion ansatz. Additionally, throughout this section  $\tau$  is referring to a new time variable and is unrelated to the original timescale, used for the discrete problem, early in Section 2.)

At the  $\mathcal{O}(1)$  level, the system (9)–(13) reduces to the Wright-Fisher linear diffusion equation ([13], [36], [24])

$$\begin{aligned} \partial_\tau \phi_0(p, \tau, T) - \mathcal{L}\phi_0(p, \tau, T) &:= \partial_\tau \phi_0(p, \tau, T) \\ -\frac{1}{2}\partial_{pp}(p(1-p)\phi_0(p, \tau, T)) &= 0, \end{aligned} \tag{19}$$

where  $\mathcal{L}$  will be referred to as the Wright-Fisher operator, with the constraint

$$\lim_{p \rightarrow 0^+} p\phi_0 = \lim_{p \rightarrow 1^-} (1-p)\phi_0 = 0. \tag{20}$$

We assume that  $p(1-p)\phi_0^2$  is integrable on  $[0, 1]$ , a condition satisfied by all continuous functions that satisfy (20) (and hence by all biologically reasonable solutions of 19). The solution of equation (19) with the constraint (20) was described by Kimura [24]. Specifically, the operator  $\mathcal{L}$  has eigenvalues  $\lambda_i = (i + 1)(i + 2)/2$  ( $i = 0, 1, \dots$ ) and eigenfunctions  $\psi_i$  which are scaled and translated Gegenbauer polynomials, which form an orthonormal set in  $L^2([0, 1])$  with respect to the weight  $w(p) = 6p(1-p)$  [27]. The solution to (19) can therefore be written as a linear

combination of Wright-Fisher eigenfunctions with coefficients that vary slowly in time:

$$\phi_0(p, \tau, T) = \sum_{i=0}^{\infty} a_i(T) e^{-\lambda_i \tau} \psi_i(p). \tag{21}$$

A standard procedure (see for example [23]) can then be used to find the values of the coefficients  $a_i(T)$ .

Keeping just the first  $m + 1$  Gegenbauer polynomials  $\psi_0 \dots \psi_m$  in the expansion (21) for  $\phi_0$ , approximations were obtained for  $m = 2, 4$ , and  $6$  with the aid of the symbolic mathematics package Maple (Maple 9 worksheets available upon request). If  $m = 2$ , the expansions involve projecting onto the first three eigenfunctions:

$$\psi_0(p) = 1, \quad \psi_1(p) = 2\sqrt{5}(p - \frac{1}{2}), \quad \psi_2(p) = \sqrt{14}(\frac{1}{4} - 5(p - \frac{1}{2})^2).$$

For example, at  $\mathcal{O}(1)$ , in the case where no loci are initially fixed we have for  $m = 2$

$$\phi_0(p, t) = a_{00}e^{-t} + a_{10}e^{-3t}\psi_1(p) + a_{20}e^{-6t}\psi_2(p) \tag{22}$$

$$\bar{z}_0(t) = \frac{\sqrt{5}\alpha}{3}a_{10} \quad S_0^2(t) = \frac{2\alpha^2}{3}a_{00}e^{-t} \tag{23}$$

for constants  $a_{00}, a_{10}, a_{20}$  determined by the initial data  $\phi(p, 0)$ . Also, above we provide the  $\mathcal{O}(1)$  portion of the total additive genetic variance  $S^2$ , as computed by integrating  $\phi_0$  against  $4\alpha^2 p(1 - p)$ .

When  $t$  is large, (22) indicates that  $\phi_0$  will be dominated by a steadily decaying horizontal profile and that  $S_0^2(t)$  will decay at the same rate. Additionally, if the initial data is even about  $p = 1/2$  ( $a_{10} = 0$ ) then  $\bar{z}_0 = 0$ . This is because the trait mean is computed by integrating  $\phi_0$  (plus  $\delta$ -functions at  $p = 0$  and  $p = 1$ ) against  $(p - 1/2)$ . Since this weight is odd about  $p = 1/2$ , the integral will only pick up the even parts of  $\phi_0$ .

We note a subtlety in the interpretation of  $\bar{z}_0$ : this quantity does not depend on  $t$ . This is reasonable when one recalls the standard interpretation of the Wright-Fisher equation (19) [13], [36], [24]. In this interpretation, when  $\epsilon = 0$ ,  $\bar{z}_0$  represents the mean or expected behavior of the trait mean in the absence of selection. Although the trait mean of a particular population will in general change over time in response to random drift, the expected trait mean averaged over all possible populations will not change. In our interpretation, the fact that  $\bar{z}_0$  is constant likewise reflects the absence of stochastic effects due to genetic sampling. However, it is not clear that the full system gives an unbiased representation of the mean behavior over replicate populations, because of the nonlinearity that arises from coupling of loci through the trait mean. (In contrast to the behavior of the trait mean, genetic variance must decline even in the absence of selection, since the model does not include mutation or migration; thus  $S_0^2(t)$  does decay to 0 as  $t \rightarrow \infty$  even when  $\epsilon = 0$ .)

The coefficient of  $\epsilon$  for the trait mean is:

$$\bar{z}_1(t) = -\frac{a_{00}\alpha}{9} \left( 2\sqrt{5}\alpha a_{10} - 6z_{opt} \right) (1 - e^{-t}) \tag{24}$$

The effect of selection is apparent at the  $\mathcal{O}(\epsilon)$  level in that  $\bar{z}_1(t)$  (unlike  $\bar{z}_0$ ) is time-dependent. Furthermore, even if the initial data is even about  $p = 1/2$ ,  $\bar{z}_1$  will in general differ from 0 at times  $t > 0$ . This is because selective pressure will cause allele frequencies to shift preferentially toward one extreme ( $p = 0$  or  $p = 1$ ), breaking any symmetry about  $p = 1/2$ . Such behavior is reflected in the skewness of the plots of solutions given in Figures 1 and 2.

To check the validity of the asymptotics and to determine parameter regimes in which the asymptotics provide a useful approximation to solutions of the full system (9)–(13), asymptotic approximations for  $S^2(t)$  and for the trait mean  $\bar{z}(t)$  were compared with those obtained from numerical solutions of the full system for a single value of  $\alpha$ , several initial  $p$ -distributions, and a range of selection parameters  $V_s$  and  $z_{opt}$ . The comparisons indicate that the two-time-scale asymptotic approximations can predict future values of the trait mean and total additive genetic variance from initial data accurately for the cases of weak selection.

In addition, comparisons of the asymptotics with data for a series of runs with various initial  $x$ -distributions provided evidence that the absolute value of the difference between the asymptotically predicted and numerically computed values of  $\bar{z}(t_f)$  and  $S^2(t_f)$  decayed like  $\epsilon^2$  as  $\epsilon \rightarrow 0$ , giving  $\mathcal{O}(\epsilon)$  convergence of the asymptotics and numerical data as expected. Specifics regarding runs with initial data

$$\phi(p, 0) = 1 - 2(p - 1/2)^3$$

with  $z_{opt} = 0$  are given in Table 1. (Recall that each numerical run was stopped when 0.5% of initially unfixed loci remained unfixed. This stopping time is  $t_f$ .) Here, asymptotic estimates of  $\bar{z}(t_f)$  with  $m = 2$  failed to show  $\mathcal{O}(\epsilon)$  convergence. Using  $m = 6$ , however, gave the expected rate of convergence and matched the numerically computed  $\bar{z}(t_f)$  with relative error  $< .001$  for all but the two largest values of  $\epsilon$  tried; see Table 1.

**Table 1.** Comparison of asymptotically predicted and numerically computed values of  $\bar{z}(t_f)$  using initial data cubic in  $p$ ,  $\alpha = .5$ ,  $1/2V_s$  from  $10^1$  to  $10^{-4}$ , and  $z_{opt} = 0$ . Logarithms are base 10. Here  $abserr := |\bar{z}_{asympt}(t_f) - \bar{z}_{num}(t_f)|$ ,  $relerr := |\bar{z}_{asympt}(t_f) - \bar{z}_{num}(t_f)|/|\bar{z}_{num}(t_f)|$ , where  $\bar{z}_{num}$  and  $\bar{z}_{asympt}$  are respectively the numerically computed and asymptotically predicted values of  $\bar{z}$ .  $m$  denotes the order of the Gegenbauer polynomials used in the asymptotics.

$\epsilon$	$\log(abserr) (m = 2)$	$\log(abserr) (m = 6)$	$\log(relerr) (m = 6)$
10	-1.3739	-1.2305	1.8127
1	-2.2213	-2.9089	-1.1628
0.1	-2.1599	-4.8703	-3.2539
0.01	-2.1476	-6.865	-5.2615
0.001	-2.1463	-8.8673	-7.2650
0.0001	-2.1461	-10.861	-9.2589

## 5. Insensitivity to multiple effect sizes

In Section 3, we presented simulations of a population where all QTL have only one effect size,  $\alpha$ . In such a population, the mean of the distribution of locus effect sizes is  $\alpha$  and the variance is 0. We now present a more detailed study of response of the trait mean to selection in a population where not all QTL have the same effect size.

To do this, a series of numerical solutions to (9)–(13) was computed. Each run used an effect size distribution with two values of  $\alpha$ , as such distributions are the simplest possible distributions of  $\alpha$  with nonzero variance. We denote the two values of effect size by  $\alpha_1$  and  $\alpha_2$ , with  $\alpha_1 \leq \alpha_2$ . The fraction of loci with the high value of  $\alpha$ ,  $P(\alpha = \alpha_2)$ , was chosen from 1/2, 1/3 and 1/5.

For each set of runs, an initial distribution of allele frequencies  $p$  was specified for loci with each effect size. This initial  $p$ -distribution was always uniform for loci with the low effect size  $\alpha_1$ , i.e. we specified the conditional probability densities

$$\phi(p, \alpha_1, 0) / \int \phi(\xi, \alpha_1, 0) d\xi \equiv 1$$

for all runs. For loci with the high effect size  $\alpha_2$ , the initial  $p$ -distribution was varied among three different profiles, i.e. we specified the conditional probability densities

$$\phi(p, \alpha_2, 0) / \int \phi(\xi, \alpha_2, 0) d\xi = g(p),$$

with  $g(p)$  varying among the uniform distribution

$$g_{unif}(p) \equiv 1,$$

a quartic profile concentrated near  $p = \frac{1}{2}$ ,

$$g_{peak}(p) = \frac{21}{16} - 3(p - \frac{1}{2})^2 - 5(p - \frac{1}{2})^4$$

and a quartic profile concentrated near  $p = 0$  and  $p = 1$ ,

$$g_{valley}(p) = 80(p - \frac{1}{2})^4.$$

The parameters were  $1/2V_s = .1$  and  $z_{opt} = .1$ . A set of runs was performed, varying the mean  $\bar{\alpha} = p_1\alpha_1 + p_2\alpha_2$  and the coefficient of variation

$$cv(\alpha) = \sqrt{p_1(\alpha_1 - \bar{\alpha})^2 + p_2(\alpha_2 - \bar{\alpha})^2} / \bar{\alpha}$$

of the  $\alpha$ -distribution among the values

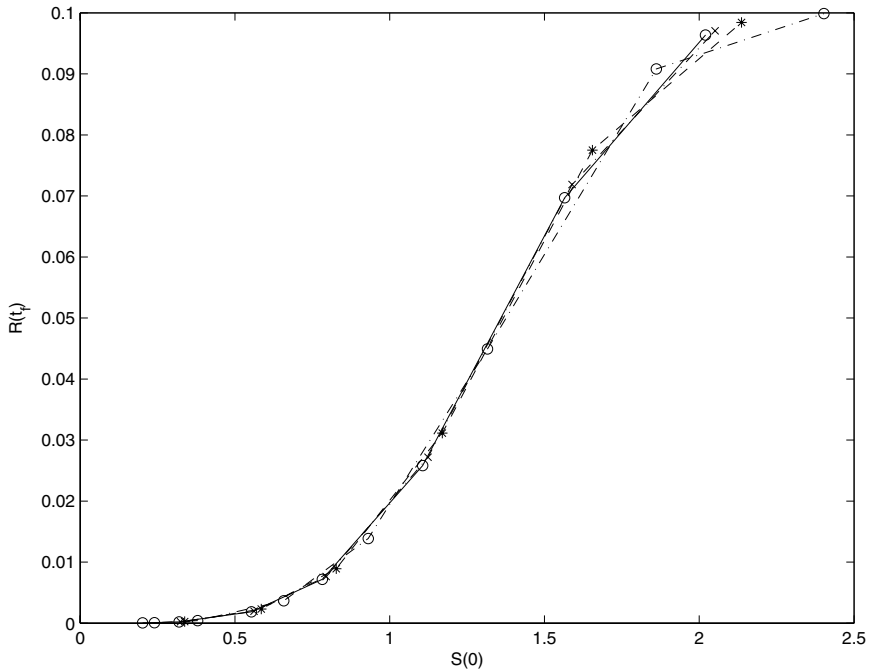
$$\bar{\alpha} \in \{.05, .125, .375, .75, 1.5, 3, 5\}, \quad cv(\alpha) \in \{0, .05, .1, .25, .5, 1\}.$$

Each pair  $(\bar{\alpha}, cv(\alpha))$  determines a pair of effect sizes,  $(\alpha_1, \alpha_2)$ , giving a total of 42 runs. Each run was continued until a time  $t = t_f$  at which only 0.5% of originally unfixed loci remained unfixed. Here we will focus on the case where the fraction of loci with the high value of  $\alpha$  is 1/2 and initial distributions of allele frequencies

$p$  are uniform for both values of  $\alpha$ . Results for other initializations were similar (data not shown).

The values of  $\bar{z}(t_f)$  for the various runs were compared with the values of the initial total additive genetic variance  $S^2(0)$ . One should expect a closer approach to the optimum trait value (i.e.  $\bar{z}(t_f)$  closer to  $z_{opt}$ ) when the initial total additive genetic variance  $S^2(0)$  is larger [12], as shown by the simulations for a single effect size in Section 3. For this initial data,  $S^2(0) = 4\bar{\alpha}^2(1 + cv(\alpha)^2)/6$ . To keep the relation between  $\bar{\alpha}$  and  $S^2(0)$  transparent, we therefore studied the relation between  $S^2(0)$  and  $\bar{z}(t_f)$  for fixed values of  $cv(\alpha)$ .

In Figure 6, we plot the final trait mean  $\bar{z}(t_f)$  versus  $S(0) := \sqrt{S^2(0)}$  for four values of  $cv(\alpha)$ . This figure suggests that the final trait mean depends on the mean and variance of the distribution of effect sizes only insofar as these quantities influence the total additive genetic variance of the initial population.



**Fig. 6.** The results of twenty-four two- $\alpha$  simulations of (9)–(13). (Results of eighteen additional simulations followed the same pattern but are omitted from the graphics for clarity.) The model parameters are  $1/2V_s = .1$  and  $z_{opt} = .1$ . The QTL effect size is  $\alpha_1$  with probability  $p_1 = 1/2$  and is  $\alpha_2$  with probability  $p_2 = 1/2$ . The initial distributions are:  $\phi(x, \alpha_i, 0) = p_i$  for  $0 < x < 1$  and  $\phi_0(\alpha_i, 0) = \phi_1(\alpha_i, 0) = 0$  for  $i = 1, 2$ . The effect sizes are chosen to achieve a given mean  $\bar{\alpha}$  and coefficient of variation  $cv(\alpha)$  and the solution is computed the time  $t_f$  when all but 0.5% of the QTL are fixed. For each value of  $cv(\alpha)$ , we then plot  $\bar{z}(t_f)$  as a function of  $S(0)$ , the square root of initial total additive genetic variance, to see how close  $\bar{z}(t_f)$  is to  $\bar{z}_{opt}$ . Solid lines with circles denote runs with  $cv(\alpha) = 0$ , dot-dashed lines with x's  $cv(\alpha) = .25$ , dashed lines with stars  $cv(\alpha) = .5$ , and dot-dashed lines with circles  $cv(\alpha) = 1$ .

## 6. Discussion

We have introduced a diffusion model (9)–(13) for the joint distribution of absolute locus effect sizes and allele frequencies for loci contributing to an additive quantitative trait in a haploid, panmictic population. It is a “mesoscale” model in that it explicitly incorporates a finite number of loci with finite (i.e. non-infinitesimal) effects, but does not track the evolution of allele frequencies at specific individual loci. The model is designed to approximate a particular discrete model exactly in the limit as both population size and the number of loci affecting the trait tend to infinity. We have studied the long-time behavior of solutions to the diffusion system, using formal multiple-timescale asymptotics, for the case when all loci have the same effect size. Finally, we have presented numerical solutions of the system for the case where loci can take on either of two distinct effect sizes, varying the proportion of loci taking on each effect size. We confined our attention in the present work to these arguably unrealistic distributions of effect sizes because our aim was to explore the influence of the mean and variance of locus effect size on the response of the population trait (phenotypic) mean to selection, and distributions with two effect sizes are the simplest distributions with nonzero variance.

The long-time behavior of solutions to the diffusion system can be studied in the case of weak selection by using the case of no selection as a base case. When no selection is present, the diffusion system (9)–(13) reduces in form, though not in interpretation, to the linear Wright-Fisher equation (19) [13], [36]. Kimura [24] showed that the solution to (19) is an infinite sum of decaying eigenmodes, i.e. of allele frequency distributions each going to fixation at a characteristic rate. The formal asymptotics presented here indicate that weak selection alters this picture by changing (or modulating) the rates at which the various modes decay, on a slow time scale determined by the strength of selection relative to random drift. Moreover, numerical simulations indicate that an approximation consisting of a linear combination of a small number of modulated decaying modes can be used to make accurate predictions of the response to selection in the full model (9)–(13). The accuracy of the predictions appears to depend on the higher moments of the allele frequency distribution at the time when the approximation is fitted to numerical data as well as on the intensity of selection; data with larger higher moments requires more terms in the approximation to obtain predictions with a fixed amount of accuracy.

Through numerical solution of the diffusion system for the case of two distinct effect sizes, we found that the slope and shape of the curve relating initial total additive genetic variance to the “final” value of the population trait mean was essentially independent of the coefficient of variation of effect size when more than five loci contributed to the trait and all model parameters, as well as the form of the initial data, were held constant. This confirms the results of [22] in suggesting that knowledge of the genetic architecture of a quantitative trait will not aid in predicting response to selection when selection is weak and mutation is negligible.

Finally we note a number of important limitations of the diffusion model (9)–(13). The model ignores mutation and is thus suitable for predicting response to selection only on a short evolutionary timescale. It also ignores dominance

and epistasis, environmental effects, gametic phase disequilibrium, pleiotropy and population structure, and (like all diffusion models) assumes that selection is weak. Thus its usefulness should lie in providing quantitative analytical predictions that can be used as a basis for comparison with simulation results (e.g. of response to selection) for models that include a broader range of genetic forces but are analytically intractable. In addition, the model ignores two kinds of (broadly defined) sampling variation. One is variation due to genetic sampling, which can be expected to produce different evolutionary outcomes in replicate populations with the same initial conditions and subject to the same evolutionary forces. In ignoring such variation the diffusion model resembles, for example, a Markov chain model for the evolution of allele frequencies at a small number of loci [21, 6.1].

The other type of variation ignored by the model is variation due to “locus sampling”, i.e. due to the fact that only a finite number of loci contribute meaningfully to any given trait. To see this, suppose that all loci have the same effect size. In this case, the solution of the diffusion model (9)–(13) is a continuous probability distribution showing the percentage of loci expected to fall within any given range of allele frequencies at any time. However, the expected frequencies would not be completely realized in actual populations because these populations have only a finite number of loci. Preliminary stochastic simulations (data not shown) indicate that a large amount of variation in response to selection (between replicate populations) can occur due to this effect when the number of loci contributing to a trait is low to moderate (say, 10), as may be expected for many though not all traits.

Locus sampling variation must be taken into account in hypothesis tests for the role of selection in creating an observed phenotypic difference between two populations; this is implicit in the tests proposed by Orr [29]. Ignoring it means that the present model would not be suitable for determining the variance of the sampling distribution of a test statistic under a specific alternative to the null hypothesis of no selection in such a test, although it might be suitable for determining the mean of the sampling distribution. Rather, to determine this type of sampling variance and related quantities it will be essential to conduct stochastic simulations of this evolution. Such simulations, ideally in conjunction with analytical studies, can determine the amount of variation between populations that is due to finite population size and finite locus number.

*Acknowledgements.* The work of JRM and MBH was partially supported by NSF grant DMS-0201173. The work of MCP was partially supported by NSERC grant number 250305-02 and by an Alfred P. Sloan fellowship. JRM thanks the University of Toronto and the University of Maryland for hospitality provided during the preparation of this work. Likewise, MCP thanks Georgetown University. JRM and MBH thank D. Hawthorne for helpful discussions. The authors thank S. Gavrillets and an anonymous reviewer for comments on an earlier version of the manuscript. Computational resources were provided by Georgetown’s Advanced Research Computing Initiative with generous assistance from A. Miles and J. Cannata.

## A. Appendix — Derivation of the Diffusion Model

Below we sketch a formal derivation of the diffusion model along standard lines [10, §4.1]. We have not undertaken a derivation from first principles, although one should be possible along the lines of [9]. Although the derivation takes a standard form, there is an important difference between this diffusion model and the classical diffusion models of population genetics: the approximation to the discrete system should be exact not in the limit of infinite population size  $N$ , but in the joint limit  $N \rightarrow \infty, n \rightarrow \infty$  where  $n$  is the number of loci with any given effect size.

We recall the discrete model (1), (2), (6) with dimensional parameters. To begin, let  $z_0$  be a characteristic trait scale, e.g. the initial phenotypic standard deviation for the trait in question. Take  $\tilde{V}_s$ , a dimensionless number, such that

$$V_s = \frac{N z_0^2 \tilde{V}_s}{n} \quad (\text{A1})$$

where  $n$  is the number of loci and  $N$  is the population size.

Let  $\tilde{\alpha}$  be the dimensionless number such that

$$\alpha = \frac{\tilde{\alpha} z_0}{n}. \quad (\text{A2})$$

Similarly, let  $\tilde{z}$  and  $\tilde{z}_{opt}$  be dimensionless numbers such that

$$\bar{z} = \tilde{z} z_0, \quad z_{opt} = \tilde{z}_{opt} z_0.$$

Finally, rescale time so that that one unit of time comprises  $N$  generations:  $t = \tau/N$ .

The reasoning for how  $\tilde{V}_s$ ,  $\tilde{\alpha}$ , and  $t$  are assumed to depend on  $n$  and  $N$  will be explained below.

Using  $\Phi$  as defined before equation (7), we introduce the joint probability distribution of allele frequency  $p_j = j/N$  and effect size  $\tilde{\alpha}$  at time  $t_k = k/N$ :

$$\phi_D(p_j, \tilde{\alpha}, t_k) := \frac{1}{n} \Phi \left( p_j, \frac{\tilde{\alpha} z_0}{n}, k \right) \quad (\text{A3})$$

(The  $D$  is a reminder that we are considering a discrete stochastic process.) At each time  $t_k$ , the probability distribution satisfies

$$\sum_{\tilde{\alpha}} \sum_{j=0}^N \phi_D(p_j, \tilde{\alpha}, t_k) = 1. \quad (\text{A4})$$

Similarly, we define the conditional probability  $\phi_D(p_i, t_l | p_j, \tilde{\alpha}, t_k)$ , which is the probability that a locus with effect size  $\tilde{\alpha}$  that has frequency  $p_j$  at time  $t_k$  will have frequency  $p_i$  at time  $t_l$ . (Note that the symbol  $\phi_D$  is doing double duty.)

Let  $\delta t = 1/N$  represent one generation. The Chapman-Kolmogorov equation states that if  $t_k \leq t_l \leq t_l + \delta t$  then

$$\phi_D(p_i, t_l + \delta t | p_j, \tilde{\alpha}, t_k) = \sum_{m=0}^N \phi_D(p_i, t_l + \delta t | p_m, \tilde{\alpha}, t_l) \phi_D(p_m, t_l | p_j, \tilde{\alpha}, t_k).$$

Multiplying by  $\phi_D(p_j, \tilde{\alpha}, t_k)$  and summing over  $p_j$  gives

$$\phi_D(p_i, \tilde{\alpha}, t_l + \delta t) = \sum_{m=0}^N \phi_D(p_i, t_l + \delta t | p_m, \tilde{\alpha}, t_l) \phi_D(p_m, \tilde{\alpha}, t_l). \tag{A5}$$

The domain of  $\phi_D$  is discrete:  $p_j \in \{0, 1/N, \dots, (N - 1)/N, N\}$  and  $t_k \in \{0, 1/N, 2/N \dots\}$ . Similarly, the range is discrete:  $\phi_D$  takes on values in the set  $\{0, 1/n, 2/n, \dots, (n - 1)/n, 1\}$ . Since there are  $N + 1$  possible values of  $p_j$  and  $\sum_{\tilde{\alpha}} \sum_j \phi_D(p_j, \tilde{\alpha}, t_k) \equiv 1$ , one would expect the values taken on by  $\phi_D$  to be proportional to  $1/N$ . As  $n \rightarrow \infty$  and  $N \rightarrow \infty$ , therefore, one would hope that  $N\phi_D$  would become a good approximation of a function defined for  $p \in [0, 1]$  and continuous for  $p \in (0, 1)$  and  $t \in [0, \infty)$ .

In this direction, we introduce  $\phi(p, \tilde{\alpha}, t)$ , which is a continuous function for  $p \in (0, 1)$  and  $t \geq 0$ , and  $\delta$ -functions  $\phi_0(\tilde{\alpha}, t)\delta_0(p)$  and  $\phi_1(\tilde{\alpha}, t)\delta_1(p)$ . The delta functions represent the loci that have effect size  $\tilde{\alpha}$  and allele frequency 0 or 1 (reflecting fixation) at time  $t$ . It is expected that  $\phi + \phi_0\delta_0 + \phi_1\delta_1$  is the density function corresponding to a diffusion process that approximates the original, discrete process (with density  $\phi_D$ ) with an error that tends to 0 as  $N$  and  $n$  go to infinity. For example, since  $\phi_D$  satisfies (A4) it's expected that  $\phi$ ,  $\phi_0$ , and  $\phi_1$  will satisfy

$$\sum_{\tilde{\alpha}} \left[ \int_0^1 \phi(p, \tilde{\alpha}, t) dp + \phi_0(\tilde{\alpha}, t) + \phi_1(\tilde{\alpha}, t) \right] = 1. \tag{A6}$$

at each time  $t$ . As we do not prove such a diffusion process exists, from here on all steps in the derivation are formal.

To find the PDE satisfied by the continuous density  $\phi$ , we start with the continuous version of the Chapman-Kolmogorov equation (A5):

$$\phi(p, \tilde{\alpha}, t + \delta t) = \int_0^1 \phi(p, t + \delta t | \tilde{p}, \tilde{\alpha}, t) \phi(\tilde{p}, \tilde{\alpha}, t) d\tilde{p}. \tag{A7}$$

Let  $f$  be a smooth test function that vanishes at  $p = 0$  and  $p = 1$ . Multiplying (A7) by  $f(p)$  and integrating gives

$$\int f(p) \phi(p, \tilde{\alpha}, t + \delta t) dp = \iint f(p) \phi(p, t + \delta t | \tilde{p}, \tilde{\alpha}, t) \phi(\tilde{p}, \tilde{\alpha}, t) d\tilde{p} dp. \tag{A8}$$

Letting  $\delta p = p - \tilde{p}$ , consider the Taylor expansion of  $f$ :

$$f(p) = f(\tilde{p}) + \delta p f'(\tilde{p}) + \frac{1}{2}(\delta p)^2 f''(\tilde{p}) + \dots$$

Thus (A8) can be written as

$$\begin{aligned} \int f(p)\phi(p, \tilde{\alpha}, t + \delta t) dp &= \iint f(\tilde{p})\phi(\tilde{p}, \tilde{\alpha}, t)\phi(p, t + \delta t|\tilde{p}, \tilde{\alpha}, t) d\tilde{p} dp \\ &+ \iint f'(\tilde{p})\phi(\tilde{p}, \tilde{\alpha}, t) [(\delta p)\phi(p, t + \delta t|\tilde{p}, \tilde{\alpha}, t)] d\tilde{p} dp \\ &+ \frac{1}{2} \iint f''(\tilde{p})\phi(\tilde{p}, \tilde{\alpha}, t) [(\delta p)^2\phi(p, t + \delta t|\tilde{p}, \tilde{\alpha}, t)] d\tilde{p} dp \\ &+ \sum_{k=3}^{\infty} \frac{1}{k!} \iint \frac{\partial^k f}{\partial p^k}(\tilde{p}) \phi(\tilde{p}, \tilde{\alpha}, t) [(\delta p)^k\phi(p, t + \delta t|\tilde{p}, \tilde{\alpha}, t)] d\tilde{p} dp. \end{aligned} \tag{A9}$$

Integration with respect to  $p$  on the right-hand side of (A9) gives

$$\begin{aligned} &\int f(p)\phi(p, \tilde{\alpha}, t + \delta t) dp - \int f(\tilde{p})\phi(\tilde{p}, \tilde{\alpha}, t) d\tilde{p} \\ &= \int f'(\tilde{p})\phi(\tilde{p}, \tilde{\alpha}, t)E [(\delta p)|\tilde{p}, \tilde{\alpha}, t] d\tilde{p} \\ &+ \frac{1}{2} \int f''(\tilde{p})\phi(\tilde{p}, \tilde{\alpha}, t)E [(\delta p)^2|\tilde{p}, \tilde{\alpha}, t] d\tilde{p} \\ &+ \sum_{k=3}^{\infty} \frac{1}{k!} \int \frac{\partial^k f}{\partial p^k}(\tilde{p}) \phi(\tilde{p}, \tilde{\alpha}, t)E [(\delta p)^k|\tilde{p}, \tilde{\alpha}, t] d\tilde{p}. \end{aligned}$$

The next step is to change the (dummy) variable of integration to  $p$  through-out, to divide both sides by  $\delta t$ , and to take the  $\delta t \rightarrow 0$  limit, transforming the above into the weak formulation of an evolution equation for  $\phi$ . To do this, we need to calculate the expected values above and determine how they depend on  $n$  and  $N$ .

Above,  $\delta p$  denotes the change in allele frequency over a time interval  $\delta t$  for a locus with effect size  $\tilde{\alpha}$  and allele frequency  $p$  at time  $t$ . At the discrete level, this is determined by a Bernoulli process with probability  $p'(p_j, \tilde{\alpha}, t_k)$  given by (6). Using the moments for the discrete Bernoulli process, we compute the expected values:

$$\begin{aligned} E [\delta p|p_j, \tilde{\alpha}, t_k] &= p'(p_j, \tilde{\alpha}, t_k) - p_j \\ E [(\delta p)^2|p_j, \tilde{\alpha}, t_k] &= \frac{p'(p_j, \tilde{\alpha}, t_k)[1 - p'(p_j, \tilde{\alpha}, t_k)]}{N} + (p'(p_j, \tilde{\alpha}, t_k) - p_j)^2 \\ E [(\delta p)^3|p_j, \tilde{\alpha}, t_k] &= \frac{p'(1 - p')(1 - 2p')}{N^2} + \frac{3p'(1 - p')(p' - p_j)}{N} + (p' - p_j)^3. \end{aligned}$$

Our expectation is that  $\phi$  will be the density of a diffusion process whose moments are approached by those of the Bernoulli process as  $n \rightarrow \infty$  and  $N \rightarrow \infty$ . Accordingly we use (A1) and (A2) and expand the above expected values in powers of  $1/N$ . If we assume that  $n/N \rightarrow 0$  as  $N, n \rightarrow \infty$  (which is reasonable as typically

the number of loci contributing to a trait is much less than the population size) then this yields

$$\begin{aligned}
 E[(\delta p)|p, \tilde{\alpha}, t] &= \frac{2\tilde{\alpha}p(1-p)(\tilde{z}_{opt} - \tilde{z}(t))}{\tilde{V}_s N} + o\left(\frac{1}{N}\right) \\
 E[(\delta p)^2|p, \tilde{\alpha}, t] &= \frac{p(1-p)}{N} + o\left(\frac{1}{N}\right) \\
 E[(\delta p)^3|p, \tilde{\alpha}, t] &= o\left(\frac{1}{N}\right).
 \end{aligned}$$

We expect  $E[(\delta p)^k|p, \tilde{\alpha}, t]$  to be  $o(1/N)$  for all  $k \geq 4$ . It is here that the scalings for  $V_s$  and  $\alpha$  were key. That effect size  $\alpha$  should scale like  $1/n$  was determined by other concerns (see below). Given this scaling for  $\alpha$ , if one seeks a scaling for  $V_s$  that would result in  $E[(\delta p)]$  and  $E[(\delta p)^2]$  having nontrivial contributions but the higher moments vanishing, this determines the scaling for  $V_s$ . We now take the  $n, N \rightarrow \infty$  limit and find that  $\phi$  satisfies

$$\phi_t = -(M\phi)_p + \frac{1}{2}(V\phi)_{pp} \tag{A10}$$

in the sense of distributions, where

$$M = 2\frac{\tilde{\alpha}}{\tilde{V}_s} p(1-p)(\tilde{z}_{opt} - \tilde{z}(t)), \tag{A11}$$

$$V = p(1-p). \tag{A12}$$

Since the test function  $f$  vanishes at  $p = 0, 1$  if  $\phi$  is smooth then the PDE hold in the classical sense for  $p \in (0, 1)$  and  $t \geq 0$ .

To complete the formal derivation of the diffusion system, we express the trait mean  $\tilde{z}(t)$  in terms of  $\phi(p, \tilde{\alpha}, t)$ ,  $\phi_0(\tilde{\alpha}, t)$  and  $\phi_1(\tilde{\alpha}, t)$ . It follows from (A6) and (A10) that the fractions of the loci that have effect size  $\tilde{\alpha}$  and are fixed at  $p = 0$  or  $p = 1$  equal

$$\phi_0(\tilde{\alpha}, t) = \phi_0(\tilde{\alpha}, 0) - \int_0^t \sum_{\tilde{\alpha}} \lim_{p \rightarrow 0^+} \left( [M\phi(p, \tilde{\alpha}, s) - \frac{1}{2}\partial_p(V\phi(p, \tilde{\alpha}, s))] \right) ds$$

$$\phi_1(\tilde{\alpha}, t) = \phi_1(\tilde{\alpha}, 0) + \int_0^t \sum_{\tilde{\alpha}} \lim_{p \rightarrow 1^-} \left( [M\phi(p, \tilde{\alpha}, s) - \frac{1}{2}\partial_p(V\phi(p, \tilde{\alpha}, s))] \right) ds$$

and hence the total contributions to the trait mean  $\tilde{z}(t)$  from loci fixed at 0 and 1 respectively are

$$\begin{aligned}
 \tilde{z}_0(t) &= 2 \sum_{\tilde{\alpha}} \tilde{\alpha} \left(0 - \frac{1}{2}\right) \phi_0(\tilde{\alpha}, t), \\
 \tilde{z}_1(t) &= 2 \sum_{\tilde{\alpha}} \tilde{\alpha} \left(1 - \frac{1}{2}\right) \phi_1(\tilde{\alpha}, t).
 \end{aligned} \tag{A13}$$

To obtain the contribution to the trait mean from unfixed loci, we recall the discrete case, for which (1) and (7) apply:

$$\begin{aligned} \tilde{z}_D(t_k) &= \sum_{i=1}^n (2p_i - 1) \frac{\tilde{\alpha}}{n} \\ &= \sum_{\tilde{\alpha}} \sum_{j=0}^N \left(2\frac{j}{N} - 1\right) \frac{\tilde{\alpha}}{n} \Phi\left(\frac{j}{N}, \frac{\tilde{\alpha}z_0}{n}, k\right). \end{aligned}$$

Recalling  $\phi_D$ , the joint probability distribution (A3),

$$\tilde{z}_D(t_k) = \sum_{\tilde{\alpha}} \sum_{j=0}^N (2p_j - 1) \tilde{\alpha} \phi_D(p_j, \alpha, t_k).$$

(In this step, it was important to take dimensional effect size to be  $\mathcal{O}(1/n)$ .) In the derivation of the PDE, we (formally) assumed that  $N\phi_D(p, \alpha, t)$  has a limit  $\phi(p, \alpha, t)$  at all  $p$  except  $p = 0, 1$ . At  $p = 0$  and  $p = 1$ , this limit should be captured by the  $\delta$ -function weight  $\phi_0(\alpha, t)$  and  $\phi_1(\alpha, t)$  respectively. And so we consider

$$\sum_{\tilde{\alpha}} \sum_{j=1}^{N-1} (2p_j - 1) \tilde{\alpha} [N\phi_D(p_j, \tilde{\alpha}, t_k)] \left(\frac{1}{N}\right).$$

Since  $\delta p = 1/N$ , this expression has the form of a Riemann sum. Taking  $N$  to infinity, we have the trait mean for the continuous case:

$$\tilde{z}(t) = \sum_{\tilde{\alpha}} \int_0^1 (2p - 1) \tilde{\alpha} \phi(p, \tilde{\alpha}, t) dp + \tilde{z}_0(t) + \tilde{z}_1(t)$$

where  $\tilde{z}_0$  and  $\tilde{z}_1$  are defined in (A13).

This completes the formal derivation of the model.

**B. Appendix — equivalence of definitions (3) and (4) for weak selection**

We now show that in the limit of weak selection, the probability  $p'_i$  based on the mean fitness of individuals with allele  $A_i$  (see (3)) equals the probability based on the fitness of an individual with the mean phenotype  $\bar{z}_{A_i}$  (see (4)). For ease of notation, we assume  $i = 1$ .

There are  $Np_1$  individuals with allele  $A_1$ . Let  $p_{ji}$  be 1 if the  $j$ th individual has allele  $A_i$  and 0 if they have allele  $a_i$ . Then their phenotype is  $z_j := \sum_i (2p_{ji} - 1)\alpha_i$ . With this notation, the mean fitness of individuals with allele  $A_1$  is

$$w_{A_1} = \frac{1}{Np_1} \sum_{j=1}^{Np_1} w \left( \alpha_1 + \sum_{i=2}^n (2p_{ji} - 1) \alpha_i \right) =: \frac{1}{Np_1} \sum_{j=1}^{Np_1} w(\alpha_1 + \hat{z}_j)$$

where  $\hat{z}_j$  is the contribution to the  $j$ th individual's phenotype from all loci except locus 1. Similarly,

$$w_{a_1} = \frac{1}{N(1 - p_1)} \sum_{j=Np_1+1}^N w(-\alpha_1 + \hat{z}_j).$$

For any  $\hat{z}$ , let  $\mathcal{I}_{\hat{z}}$  be the set of indices such that  $j \in \mathcal{I}_{\hat{z}}$  iff  $\hat{z}_j = \hat{z}$ . Then

$$w_{A_1} = \sum_{\hat{z}} p_{\hat{z},A} w(\alpha_1 + \hat{z}) \quad \text{and} \quad w_{a_1} = \sum_{\hat{z}} p_{\hat{z},a} w(-\alpha_1 + \hat{z})$$

where

$$p_{\hat{z},A} := \frac{|\mathcal{I}_{\hat{z}} \cap \{1, \dots, Np_1\}|}{Np_1} \quad \text{and} \quad p_{\hat{z},a} := \frac{|\mathcal{I}_{\hat{z}} \cap \{Np_1 + 1, \dots, N\}|}{N(1 - p_1)}.$$

Since all pairs of loci are in gametic phase equilibrium,  $p_{\hat{z},A} = p_{\hat{z},a}$  and so we drop the  $A$  and  $a$  indices. In this notation, the mean phenotype of individuals with the  $A_1$  ( $a_1$ ) allele is

$$\bar{z}_{A_1} = \alpha_1 + \sum_{\hat{z}} p_{\hat{z}} \hat{z} \quad \text{and} \quad \bar{z}_{a_1} = -\alpha_1 + \sum_{\hat{z}} p_{\hat{z}} \hat{z}.$$

The proportion of  $A_1$  alleles in the next generation,  $p'_1$ , is defined in (3). We denote this by  $p'_{exact}$  and let  $p'_{approx}$  denote the formula given in (4):

$$p'_{exact} = \frac{p_1 w_{A_1}}{p_1 w_{A_1} + (1 - p_1) w_{a_1}} \quad p'_{approx} = \frac{p_1 w(\bar{z}_{A_1})}{p_1 w(\bar{z}_{A_1}) + (1 - p_1) w(\bar{z}_{a_1})}.$$

For large  $V_s$ , the fitness function,  $w(z) = \exp(-(z - z_{opt})^2/2V_s)$ , is approximately  $1 - (z - z_{opt})^2/2V_s$ . Using perseverance and  $\sum p_{\hat{z}} = 1$  one finds that to  $\mathcal{O}(V_s^{-2})$ ,

$$\begin{aligned} p'_{exact} &= p_1 + \frac{p_1(1 - p_1)}{8V_s} \left( \left[ \sum_{\hat{z}} p_{\hat{z}} (-2\alpha_1 + 2(\hat{z} - \hat{z}_{opt}))^2 \right] \right. \\ &\quad \left. - \left[ \sum_{\hat{z}} p_{\hat{z}} (2\alpha_1 + 2(\hat{z} - \hat{z}_{opt}))^2 \right] \right) \\ &= p_1 + 2p_1(1 - p_1)\alpha_1 \sum_{\hat{z}} p_{\hat{z}} (\hat{z}_{opt} - \hat{z}) / V_s \end{aligned}$$

and

$$p'_{approx} = p_1 + 2p_1(1 - p_1)\alpha_1 \left( \hat{z}_{opt} - \sum_{\hat{z}} p_{\hat{z}} \hat{z} \right) / V_s.$$

Thus  $p'_{exact} - p'_{approx} = \mathcal{O}(V_s^{-2})$ , as desired.

## References

1. Barton, N.H.: The maintenance of polygenic variation through a balance between mutation and stabilizing selection. *Genet. Res.* **47**, 209–216 (1986)
2. Barton, N.H., Turelli, M.: Natural and sexual selection on many loci. *Genetics* **127**, 229–255 (1991)
3. Bulmer, M.G.: The effect of selection on genetic variability. *Am. Nat.* **105**, 201–211 (1971)
4. Bürger, R.: The mathematical theory of selection, recombination, and mutation. Wiley, Chichester, 2000.
5. Bürger, R. Gimelfarb, A.: Genetic variation maintained in multilocus models of additive quantitative traits under stabilizing selection. *Genetics* **152**, 807–820 (1999)
6. Chevalet, C.: control of genetic drift in selected populations.
7. Chevalet, C.: An approximate theory of selection assuming a finite number of quantitative trait loci. *Genetics Selection Evolution* **26**, 379–400 (1994)
8. Crow, J.F., Kimura, M.: An introduction to population genetics theory. Harper and Row, New York, 1970
9. Ethier, S.N., Kurtz, T.G.: Markov processes : characterization and convergence. Wiley, New York, 1986
10. Ewens, W.J.: Mathematical population genetics. Springer, New York, 2nd edition, 2004
11. Falconer, D.S., Mackay, T.F.C.: Introduction to quantitative genetics. Longman, Essex, England, 4th edition, 1996
12. Fisher, R.A.: The correlation between relatives on the supposition of mendelian inheritance. *Trans. Royal Soc. Edinburgh* **52**, 399–433 (1918)
13. Fisher, R.A.: On the dominance ratio. *Proc. Roy. Soc. Edin.* **42**, 321–341 (1922)
14. Gavrillets, S.: Equilibria in an epistatic viability model under arbitrary strength of selection. *Journal of Mathematical Biology* **31**, 397–410 (1993)
15. Gavrillets, S., Hastings, A.: Maintenance of genetic variability under strong stabilizing selection: a two-locus model. *Genetics* **134**, 377–386 (1993)
16. Gavrillets, S., Hastings, A.: Dynamics of genetic variability in two-locus models of stabilizing selection. *Genetics* **138**, 519–532 (1994)
17. Gavrillets, S., Hastings, A.: Maintenance of multilocus variability under strong stabilizing selection. *Journal of Mathematical Biology* **32**, 287–302 (1994)
18. Gavrillets, S., Hastings, A.: Dynamics of polygenic variability under stabilizing selection, recombination, and drift. *Genetical Research* **65**, 63–74 (1995)
19. Gimelfarb, A.: Pleiotropy and multilocus polymorphisms. *Genetics* **130**, 223–227 (1992)
20. Hartl, D.L., Clark, A.G.: Principles of population genetics. Sinauer, Sunderland MA, 3rd edition, 1997
21. Hedrick, P.W.: Genetics of populations. Jones and Bartlett, Boston, 2nd edition, 2000
22. Hill, W.G., Rasbash, J.: Models of long term artificial selection in finite population. *Genet. Res.* **48**, 41–50 (1986)
23. Kevorkian, J., Cole, J.D.: Multiple scale and singular perturbation methods, volume 114 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1996
24. Kimura, M.: Diffusion models in population genetics. *J. Applied Probability* **1**, 177–232 (1964)
25. Kimura, M., Crow, J.F.: Effect of overall phenotypic selection on genetic change at individual loci. *PNAS* **75**, 6168–6178 (1978)
26. Lande, R.: The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genet. Res.* **26**, 221–235 (1976)
27. Morse, P.M., Feshbach, H.: Methods of theoretical physics, volume 1. McGraw-Hill Book Co., Inc., New York, 1953

28. Nagylaki, T.: Rate of evolution of a quantitative character. *PNAS* **89**, 8121–8124 (1992)
29. Orr, H.A.: Testing natural selection vs. genetic drift in phenotypic evolution using quantitative trait locus data. *Genetics* **149**, 2099–2104 (1998)
30. Reeve, J.P.: Predicting long-term response to selection. *Genet. Res.* **75**, 83–94 (2000)
31. Turelli, M.: Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. *Theor. Population Biol.* **25**, 138–193 (1984)
32. Turelli, M., Barton, N.H.: Dynamics of polygenic characters under selection. *Theor. Pop. Biol.* **38**, 1–57 (1990)
33. Turelli, M., Barton, N.H.: Genetic and statistical analyses of strong selection on polygenic traits: what, me normal? *Genetics* **138**, 913–941 (1994)
34. Walsh, B., Lynch, M.: Evolution and selection of quantitative traits. online preprint, <http://nitro.biosci.arizona.edu/zbook/book.html>.
35. Wright, S.: Evolution in populations in approximate equilibrium. *Journal of Genetics* **30**, 243–256 (1935)
36. Wright, S.: The differential equation of the distribution of gene frequencies. *Proc. Nat. Acad. Sci.* **31**, 382–389 (1945)