Genet. Res., Camb. (2000), **75**, *pp.* 297–313. *With* 14 *figures.* Printed in the United Kingdom

Measuring fitness by means of balancer chromosomes

N. H. BARTON[1]* AND L. PARTRIDGE[2]
[1] *Institute of Cell, Animal and Population Biology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK*
[2] *Department of Biology, University College London, 4 Stephenson Way, London NW1 2HE, UK*

(*Received 25 March 1999; and in revised form 21 June 1999*)

Summary

We present the theoretical background to a new method for measuring genetic variation for total fitness in *Drosophila*. The method allows heterozygous effects on total fitness of whole wild-type chromosomes to be measured under normal demography with overlapping generations. The wild-type chromosomes are competed against two balancer chromosomes (B1, B2, say), providing a standard genotype B1/B2 against which variation in the fitness effects of the wild-type chromosomes can be assessed. Fitness can be assessed in two ways: (i) at equilibrium of all three chromosomes under heterozygote advantage, and (ii) during displacement of one balancer by the other. Equilibrium with all three chromosomes present will be achieved only if the wild-type homozygote is not too fit, and if the fitnesses of the three heterozygotes are not too unequal. These conditions were not satisfied for any of a sample of 12 lethal-bearing chromosomes isolated from a random-bred laboratory population of *Drosophila*. At equilibrium, genotypic frequencies show low sensitivity to changes in genotypic fitness. Furthermore, where all four genotypes are viable and fertile, supplementary information from cages with only two chromosomes present and from direct measurements of pre-adult viability are required to estimate fitnesses from frequencies. The invasion method has the advantages of a greater sensitivity and of not requiring further data to estimate fitnesses if the wild-type homozygote is fertile. However, it requires that multiple samples be taken as the invasion progresses. In a discrete generation model, generation time influences fitness estimates from this method and is difficult to estimate accurately from the data. A full age-structured model can also be applied to the data from both types of experiment. For the invasion method, this gives fitness estimates close to those from the discrete generation model.

1. Introduction

Genetic variation for fitness is a key quantity in evolutionary biology. Heritable variation for net fitness sets a limit to the rate of response to natural selection (Fisher, 1930) and is crucial for theories of sex that rely on short-term benefits (Williams, 1975; Maynard Smith, 1978; Kondrashov, 1993). The genetic correlations between fitness components are also crucial for theories of good genes sexual selection (Partridge, 1983; Charlesworth, 1987; Kirkpatrick & Ryan, 1991; Burt, 1995; Kirkpatrick & Barton, 1997) and life history evolution (Williams, 1957, 1966; Charlesworth, 1980, 1990; Partridge & Harvey, 1988; Houle, 1991; Partridge & Barton, 1993). Epistatic interactions between genes determining fitness are important for theories of the evolution of recombination (Barton, 1995; de Visser *et al.*, 1998; Charlesworth, 1998), for the maintenance of quantitative genetic variation (Barton & Turelli, 1989; Whitlock *et al.*, 1995), and genetic divergence of populations under selection and drift (Cohan *et al.*, 1989; Coyne *et al.*, 1997). Population genetic theory suggests that there should be no heritable genetic variation for fitness in a population at equilibrium under natural selection (Williams, 1975; Maynard Smith, 1978). However, such an idealized population is a useful formalism rather than a reality. If the effects of mutation (e.g. Houle *et al.*, 1996) of gene flow

* Corresponding author. e-mail: n.barton@ed.ac.uk

(Endler, 1977) and of variable selection (Charlesworth, 1987) are included in models, then significant levels of genetic variation for fitness are predicted.

Despite its importance, there have been few attempts to measure genetic variation for net fitness (Partridge, 1983; Charles, 1987; Curtsinger, 1990; Burt, 1995); most past measurements have examined only components of total fitness or homozygous effects. We have developed a method for measuring net, additive, heterozygous, fitness effects of whole chromosomes extracted from random-bred wild-type populations of *Drosophila melanogaster* (Fowler *et al*., 1997). In principle, the method could be used to measure genetic variation for individual components of fitness and the genetic correlations between them. It could also be extended to measure epistatic genetic effects on fitness. The aim of this paper is to give the theoretical background to the published method (Fowler *et al*., 1997), and to its possible extensions.

The fitness effects of an allele can be defined in terms of the average number of offspring left by its bearers in one generation (Charlesworth, 1980). Lifetime reproductive success has been assumed to provide a useful phenotypic measure of fitness under many conditions, particularly for work with wild populations (Crow, 1958; Clutton-Brock, 1988; Newton, 1989; Barrowclough & Rockwell, 1993; Burt, 1995). However, where generations overlap, so that the population is age-structured, and where population size changes, the timing of offspring production becomes important in the determination of fitness (Cole, 1954; Lewontin, 1965). A measure of the intrinsic rate of increase, r, is then appropriate (Charlesworth, 1980). If lifetime reproductive success is used as a surrogate for fitness, standard measures of resemblance among relatives for lifetime reproductive success could be used to deduce the genetic variance for fitness and the contribution of additive genetic variance to the total (Price & Schluter, 1991; Houle, 1992; Burt, 1995; Rowe & Houle, 1996). Few such estimates have been attempted, and they have in general yielded values that do not differ significantly from zero (Burt, 1995).

Measurements on wild populations have the advantage of evolutionary and environmental realism, but suffer from a number of potential pitfalls. Resemblance between relatives for any trait can be attributable to environmental correlations. Fostering of offspring can reduce these (e.g. Gustaffson & Sutherland, 1988), but non-genetic maternal (e.g. Kirkpatrick & Lande, 1989; Bryant & Meffert, 1998; Fox & Savalli, 1998; Bertram & Strathmann, 1998) and paternal (Ellegren *et al*., 1996; Sheldon *et al*., 1997*a*) effects may remain. Sample sizes are in general prohibitively small, and the informative matings will not always be available. In addition, it is rarely possible to measure lifetime reproductive success of relatives over even two generations, because of the difficulties of marking individuals and of keeping track of them and all their progeny over a generation. Although accurate estimates of additive genetic variation for total fitness in nature are highly desirable, they are likely to be extremely difficult to obtain. Measurement of a subset of fitness components has therefore often been used as a surrogate for fitness (e.g. Norris, 1993; Petrie, 1994; Sheldon *et al*., 1997*b*). However, in view of the frequent finding of negative genetic correlations between different components of fitness (e.g. Rose, 1984; Zwaan *et al*., 1995; Rice, 1998), estimates of genetic variation for total fitness based on individual components are likely to be inflated.

Most attempts to measure genetic variation for fitness have been made under controlled laboratory conditions. An obvious difficulty with these studies is that the fitness of a genotype depends upon the environment in which measurements are made, with variation between genotypes possibly increasing with the harshness of the environment (Kondrashov & Houle, 1994). Standardized laboratory environments will fail to include some of the challenges of natural environments, such as predators, pathogens, competitors and physical stresses. Important sources of genetic variance for fitness such as temporal variability in selection pressures and gene flow from other populations will also be absent or reduced. Controlled laboratory studies are therefore likely to yield estimates of genetic variation for fitness that are lower than those that would be recorded in nature. When a population is moved from nature to the laboratory, adaptation to laboratory conditions often occurs (e.g. Latter & Mulley, 1995) and could lead to inflated measures of genetic variation for fitness while adaptation is in progress. Given that conditions in nature cannot be exactly replicated in the laboratory, the next best approach is therefore to work with a laboratory-adapted population; genetic variation for fitness attributable to mutation and various forms of balancing selection can then be estimated. The chromosome balance technique that we have devised (Fowler *et al*., 1997) is designed to be applied to such laboratory-adapted populations. It can be used for populations at their carrying capacity and in demographic equilibrium, in the culture conditions to which they are adapted.

A lower bound to genetic variation for fitness can be deduced from studies of the effects of mutation (Burt, 1995). For a population to maintain fitness in the face of mutation pressure it must have fitness variance at least equal to that introduced by spontaneous mutation in each generation. Several laboratory studies of *Drosophila melanogaster* have examined the effect on various fitness components of accumulation of spontaneous mutations (e.g. Mukai

et al., 1972; Houle *et al*., 1996; Shabalina *et al*., 1997), and have yielded values of about 2% per individual per generation. The approaches used in these studies have been criticized on various technical grounds, all of which could have led to an overestimate of mutational effects (Keightley, 1996). In addition, measurements of genetic variation for fitness generated by mutation in another model organism, the nematode *Caenorhabditis elegans*, led to a much lower estimate (Keightley & Caballero, 1997). A more recent study led to a somewhat higher estimate (Vassilieva & Lynch, 1999). Studies of the effects of standing genetic variance in other organisms have yielded estimates for the genomic deleterious mutation rate of 0·2–2·0% per individual per generation. These studies have the potential to provide a useful lower bound to values for the genetic variation for total fitness generated by mutation, but further studies are needed to provide estimates that can be accepted with confidence (Lynch & Walsh, 1998).

Direct estimates of genetic variation for fitness encounter the problem that a proper measure of fitness, the intrinsic rate of increase, r, cannot be made on individuals. One way around this difficulty is to make the fitness measurements on clones, which allow repeated measures on individuals of the same genotype. Asexually reproducing organisms have proved useful material for estimates of mutational decline of fitness (e.g. Lynch, 1985; Elena & Lenski, 1997) and fitness effects of adaptation to novel environments (e.g. Mongold *et al*., 1996; Lenski *et al*., 1998). However, many organisms of interest reproduce sexually, and it is desirable to develop methods for measuring their genetic variation for fitness. Repeated measures on the fitness effects of a single genotype in an organism reproducing sexually can be made by techniques that allow the same genotype to be generated repeatedly. For instance, Curtsinger (1990) assessed the effects of different X chromosomes on the total fitness of male *Drosophila*, by use of an attached X stock, which forced the wild-type X chromosomes down the male line. Ideally, genetic effects on the fitness of both sexes should be included, which can be assessed in *Drosophila* using autosomal balancer chromosomes. These contain inversions, which abolish the effects of recombination, and they carry a dominant marker mutant that is also a recessive lethal. Their presence in a fly can thus be detected, and they can never become homozygous. Balancer chromosomes have been used as standard competitors to measure the homozygous effects on fitness of chromosomes extracted from a wild-type population; a series of wild-type chromosomes were competed against a balancer over a number of generations. The balancer homozygote is lethal, and the frequency of the wild-type homozygote relative to the heterozygote with the balancer gives a measure of the fitness of the wild-type homozygote (Sved & Ayala, 1970; Sved, 1971, 1975; Wilton & Sved, 1979). The method was also used to assess the contribution of individual components of fitness to total fitness (Mackay, 1986; Partridge *et al*., 1986). This technique is based on the assumption that variation in the heterozygous effects of the wild-type chromosome is negligible; however, it is precisely these heterozygous effects that underlie heritability for total fitness in a random-bred population.

Inbred genotypes will be homozygous for a number of deleterious, recessive alleles, and they will also be homozygous for loci that may display heterozygote advantage in the original outbred population from which the chromosomes were derived. They are therefore informative about inbreeding depression, but not about genetic variation for fitness in random-bred populations. A method for assessing heterozygous fitness effects is required. Our technique uses two different balancer chromosomes with the wild-type chromosome to be assessed. Of the six possible genotypes, the two balancer homozygotes are lethal, and the wild-type homozygote may or may not be viable and fertile (for the original study (Fowler *et al*., 1997) we worked only with lethal wild-type homozygotes, and confined the theory to them; here we extend the theory to the case where the wild-type homozygote may be viable and fertile). One of the three heterozygous genotypes, containing two balancers, is present in every culture, and acts as a yardstick against which the fitness of the two balancer/wild-type heterozygotes, and the wild-type homozygote, can be assessed. Depending upon the relative fitness of the genotype (see below), two versions of the method can be used. In the first, populations are allowed to reach equilibrium under selection, and the equilibrium genotype frequencies are measured and used to deduce the relative fitness of the wild-type chromosome. In the second, the dynamics of the invasion of a pure population of one balancer heterozygote by the other balancer are estimated. We describe each approach in turn.

2. Equilibrium between two balancers and a wild-type chromosome

We denote the two balancers by B1, B2, and the wild-type chromosome by $+$. The frequencies of these chromosomes are p_1, p_2, p_+, respectively. These are counted amongst successful gametes, or amongst zygotes, *before* the inviable balancer homozygotes have been eliminated. The frequencies of the three heterozygous genotypes are $u_{1/+}$, $u_{2/+}$, $u_{1/2}$; the wild-type homozygote has frequency $u_{+/+}$. The fitnesses are denoted $W_{1/+}$, $W_{2/+}$, $W_{1/2}$, $W_{+/+}$. All fitnesses can be measured relative to the standard genotype B1/B2, which is present in all cages. Thus, we let $W_{1/2} = 1$

https://doi.org/10.1017/S0016672399004346 Published online by Cambridge University Press

throughout. Since the balancer chromosomes carry deleterious markers, we expect that $W_{1/+} > 1$, $W_{2/+} > 1$. The wild-type homozygote may be more or less fit than B1/B2.

Provided that mating is random, and provided there are no non-multiplicative interactions between male and female effects on fertility, zygotes begin in Hardy–Weinberg proportions. Thus, zygote genotypes are in the ratio $u_{1/+}:u_{2/+}:u_{1/2}:u_{+/+} = 2p_1p_+:2p_2p_+:2p_1p_2:p_+^2$. The recursions for the chromosome frequencies are:

$$p_+^* = p_+\frac{(p_1W_{1/+}+p_2W_{2/+}+p_+W_{+/+})}{\bar{W}},$$

$$p_1^* = p_1\frac{(p_2W_{1/2}+p_+W_{1/+})}{\bar{W}},$$

$$p_2^* = p_2\frac{(p_1W_{1/2}+p_+W_{2/+})}{\bar{W}}, \tag{1}$$

where

$$\bar{W} = 2p_1p_+W_{1/+}+2p_2p_+W_{2/+}+2p_1p_2W_{1/2}+p_+^2W_{+/+}.$$

This extends equation 1 of Fowler *et al.* (1997) to allow for viable and fertile $+/+$ homozygotes.

Provided that the heterozygote fitnesses are not too dissimilar, and provided the wild-type homozygote is not fit enough to fix in the population, a stable polymorphism is maintained by heterozygote advantage. If an equilibrium for all three chromosomes can be realized, it is at:

$$p_+ = \frac{(W_{1/+}+W_{2/+}-W_{1/2})\,W_{1/2}}{\{(W_{1/+}+W_{2/+}+W_{1/2})^2-2(W_{1/+}^2+W_{2/+}^2+W_{1/2}^2)-2W_{+/+}W_{1/2}\}},$$

$$p_1 = \frac{(W_{1/+}+W_{1/2}-W_{2/+})\,W_{2/+}-W_{+/+}W_{1/2}}{\{(W_{1/+}+W_{2/+}+W_{1/2})^2-2(W_{1/+}^2+W_{2/+}^2+W_{1/2}^2)-2W_{+/+}W_{1/2}\}},$$

$$p_2 = \frac{(W_{2/+}+W_{1/2}-W_{1/+})\,W_{1/+}-W_{+/+}W_{1/2}}{\{(W_{1/+}+W_{2/+}+W_{1/2})^2-2(W_{1/+}^2+W_{2/+}^2+W_{1/2}^2)-2W_{+/+}W_{1/2}\}}. \tag{2}$$

The mean fitness at this equilibrium is:

$$\bar{W} = \frac{2W_{1/+}W_{2/+}W_{1/2}-W_{+/+}W_{1/2}^2}{\{(W_{1/+}+W_{2/+}+W_{1/2})^2-2(W_{1/+}^2+W_{2/+}^2+W_{1/2}^2)-2W_{+/+}W_{1/2}\}}. \tag{3}$$

In order for this equilibrium to be feasible, all the chromosome frequencies must be positive. This requires:

$$W_{1/+}+W_{2/+} > W_{1/2},$$

$$W_{1/+}+W_{1/2} > W_{2/+}+\frac{W_{+/+}W_{1/2}}{W_{2/+}} \tag{4}$$

for an interior equilibrium to exist,

$$W_{2/+}+W_{1/2} > W_{1/+}+\frac{W_{+/+}W_{1/2}}{W_{1/+}}.$$

No equilibrium will occur if either the wild-type homozygote is very fit, or if the fitnesses of the other three genotypes are too unequal.

The approach to equilibrium involves two timescales. Suppose that $+/+$ are lethal, and the population is almost fixed for one genotype (B1/+, say). Then, after a single round of mating, the two rare genotypes must approach equal frequencies, since any mating between two different heterozygotes produces equal proportions of the three heterozygous genotypes in the next generation. This rapid approach to equal frequencies of the rare genotypes is seen in the upper part of Fig. 1, which plots genotype frequencies on a triangle. Populations quickly approach three lines which connect the vertices with the interior equilibrium. The slower increase of the rare chromosome along these trajectories is shown in the lower part of Fig. 1.

Equilibria also exist with just two chromosomes: the population may be fixed for B1/B2; it may be polymorphic for B1/+ and +/+ with

$$p_+ = \frac{W_{1/+}}{2W_{1/+}-W_{+/+}}, \quad \bar{W} = \frac{W_{1/+}^2}{2W_{1/+}-W_{+/+}};$$

or it may be polymorphic for B2/+ and +/+ with

$$p_+ = \frac{W_{2/+}}{2W_{2/+}-W_{+/+}}, \quad \bar{W} = \frac{W_{2/+}^2}{2W_{2/+}-W_{+/+}}.$$

The conditions for the edge equilibria to exist, and to be stable towards introduction of the absent chromosome, are:

$$W_{1/+}(W_{1/+}-W_{2/+}) > W_{1/2}(W_{1/+}-W_{+/+}) > 0$$

$$\text{for } \left\{\frac{\text{B1}}{+},\frac{+}{+}\right\} \text{ to exist and be stable,}$$

$$W_{2/+}(W_{2/+}-W_{1/+}) > W_{1/2}(W_{2/+}-W_{+/+}) > 0 \tag{5}$$

$$\text{for } \left\{\frac{\text{B2}}{+},\frac{+}{+}\right\} \text{ to exist and be stable,}$$

$$2W_{1/2} > W_{1/+}+W_{2/+}$$

$$\text{for } \frac{\text{B1}}{\text{B2}} \text{ to exist and be stable.}$$

https://doi.org/10.1017/S0016672399004346 Published online by Cambridge University Press

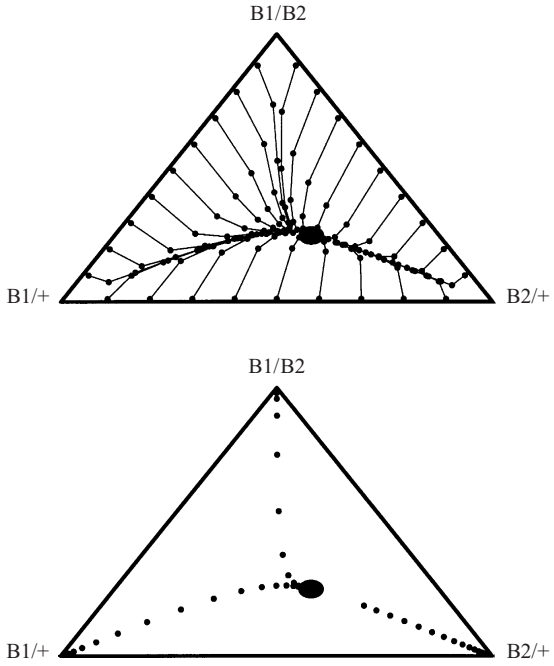

Fig. 1. Above: The frequencies of the three heterozygous genotypes (B1/+, B2/+, B1/B2) are plotted within the triangle. If the + chromosome carries a recessive lethal ($W_{+/+} = 0$), and the genotypes carrying a single balancer are fitter than B1/B2 ($W_{1/+} = 1{\cdot}5\ W_{1/2}$, $W_{2/+} = 1{\cdot}8\ W_{1/2}$), then there is a stable equilibrium at genotype frequencies $u_{1/+} = 0{\cdot}295$, $u_{2/+} = 0{\cdot}456$ (filled oval). Each line connects a trajectory, starting from a state where one of the genotypes makes up 1% of the population, and the other two are polymorphic. Points show genotype frequencies over the four following generations. Below: Fitnesses are as for the upper figure. The three sequences of points show genotype frequencies over 15 generations, starting near fixation for each of the three genotypes ($u_{1/2} = 0{\cdot}99$, etc.). Note that near fixation, the two rare genotypes approach equal frequencies after a single round of random mating.

The possible outcomes determined by conditions 4, 5) are shown in Fig. 2; in this example, $W_{+/+} = W_{1/2}/2$. Note that for some parameter values, two alternate equilibria are stable: the outcome then depends on the initial conditions. Fig. 2 also shows the fitnesses estimated by Fowler *et al.* (1997) for 12 chromosomes, each replicated twice. These chromosomes all carried recessive lethals. Fig. 2 shows that, in all but one case, one balancer would always displace the other even if the +/+ homozygote were fertile ($W_{+/+} = W_{1/2}/2$).

If the wild-type chromosome carries a recessive lethal, and if there is an interior equilibrium, then the three non-zero fitnesses can be estimated from the three chromosome frequencies, p, or equivalently, from the genotype frequencies among zygotes, u:

$$W_{1/+} = \frac{p_2(1-2p_2)}{p_+(1-2p_+)} W_{1/2}$$

$$= \frac{u_{1/2}(u_{1/+}u_{2/+} + u_{1/2}u_{1/+} - u_{1/2}u_{2/+})}{u_{1/+}(u_{1/2}u_{1/+} + u_{1/2}u_{2/+} - u_{1/+}u_{2/+})} W_{1/2},$$

$$W_{2/+} = \frac{p_1(1-2p_1)}{p_+(1-2p_+)} W_{1/2}$$

$$= \frac{u_{1/2}(u_{1/+}u_{2/+} + u_{1/2}u_{2/+} - u_{1/2}u_{1/+})}{u_{2/+}(u_{1/2}u_{1/+} + u_{1/2}u_{2/+} - u_{1/+}u_{2/+})} W_{1/2}. \quad (6)$$

Fig. 3 shows the estimated fitness as a function of the genotype frequencies. Note that variation in fitness can be estimated most accurately when the fitnesses are similar to each other (central region of Fig. 3). As fitnesses become very different, conditions for an interior equilibrium to exist become restrictive (Fig.

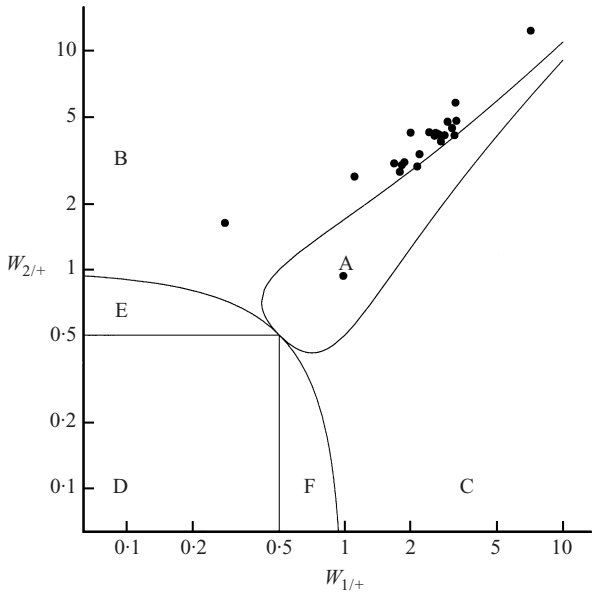

Fig. 2. Possible outcomes as a function of the fitnesses of B1/+ and B2/+, relative to B1/B2 (x and y axes, respectively). A: All four genotypes segregate. B: B2/+,+/+. C: B1/+,+/+. D: B1/B2 OR +/+ fixed. E: B1/B2 *or* B2/+,+/+. F: B1/B2 *or* B1/+,+/+. Dots show relative fitnesses of B1/+ and B2/+, estimated by Fowler *et al.* (1997). All except one lie in region B; thus, B2 invades {B1/+,+/+} to produce a population containing {B2/+,+/+}. The homozygote +/+ is assumed to have fitness 0·5 relative to B1/B2.

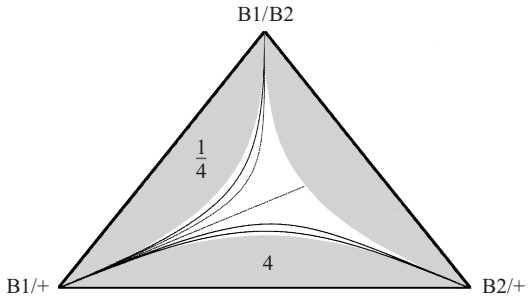

Fig. 3. The contours show the estimated fitness of B2/+ relative to B1/B2, as a function of the observed genotype frequencies. Contours are plotted for $W_{1/+}/W_{1/2} = 1/4$, 1/2, 1, 2, 4 (upper left to lower right). The estimated relative fitnesses of B1/+ would be given by the mirror image of this diagram. Equilibria cannot exist within the shaded areas.

https://doi.org/10.1017/S0016672399004346 Published online by Cambridge University Press

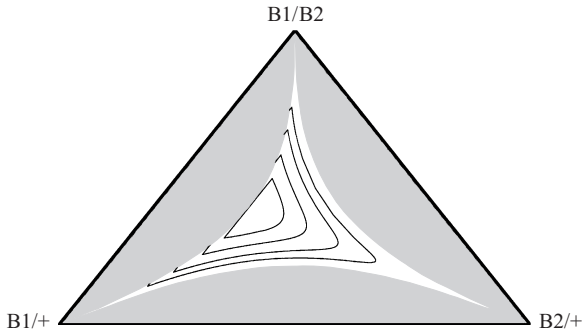

Fig. 4. The coefficient of variation (*c.v.*; i.e. standard error/mean) of estimates of $W_{1/2}$, plotted against equilibrium genotype frequencies, for a sample of 1000 zygotes. The contours show *c.v.* = 0·05, 0·1, 0·2, 0·4 (centre to edge). The accuracy of estimates of $W_{2/+}/W_{1/2}$ would be shown by the mirror image of this plot.

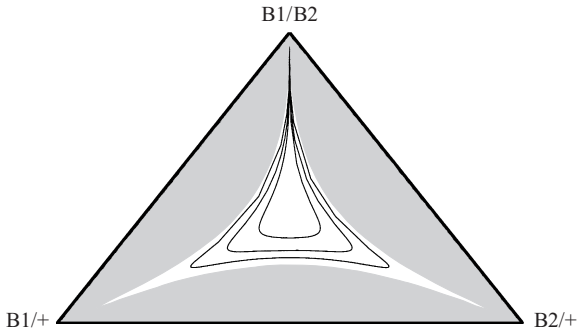

Fig. 5. The coefficient of variation (*c.v.*; i.e. standard error/mean) of estimates of $\sqrt{(W_{1/+}\,W_{2/+})}$, plotted against equilibrium difficult frequencies, for a sample of 1000 zygotes. The contours show *c.v.* = 0·1, 0·2, 0·4 (centre to edge). The greatest accuracy is when all fitnesses are equal; then *c.v.* = 0·067.

2). If such an equilibrium does exist, wide variations in the fitness of one genotype may lead to little variation in the equilibrium frequency, making estimation difficult. For example, fitnesses $W_{+/+} = 0$, $W_{1/+} = 0{\cdot}1$, $W_{2/+} = 1$, $W_{1/2} = 1$ give an equilibrium with $u_{1/2} = u_{2/+} = 0{\cdot}396$. A 10% decrease in $W_{2/+}$ makes an interior equilibrium impossible, whilst a fourfold increase in $W_{2/+}$ to 0·4 only changes the equilibrium to $u_{1/2} = u_{2/+} = 0{\cdot}381$. The poor resolution of the equilibrium method with large fitness differences is reflected by the proximity of the outer contours in Fig. 3, which show the effect of a doubling of fitness on genotype frequencies.

In practice, zygote frequencies cannot be observed directly. However, they can be inferred from frequencies among adults raised from a sample of eggs under standard conditions, provided that relative viabilities are measured by raising eggs from parents of known genotype, under the same conditions. If the wild-type homozygote is viable and fertile, then it is impossible to infer the four fitnesses from the three chromosome frequencies alone. The experiment must be supplemented by observations on populations lacking one or other balancer (Sved, 1971, 1975), which will determine the fitness of the wild-type homozygote relative to the wild-type heterozygote.

Estimates of relative fitness made using (6) will vary as a result of the binomial sampling error in the genotype frequencies $\{u_{1/+}, u_{2/+}, u_{1/2}\}$. For a large sample of n zygotes, the error variance can be found from the binomial sampling covariance of the genotype frequencies, by using the delta method. Where all fitnesses and hence frequencies are equal, the variances of the estimates of both $W_{1/+}/W_{1/2}$ and $W_{2/+}/W_{1/2}$ are $6/n$, and the covariance between these estimates is $3/n$. Fig. 4 plots the standard error of the estimate of $W_{1/+}/W_{1/2}$ against the equilibrium genotype frequencies. This estimate is most accurate when $W_{2/+}$ is extremely unfit (upper left of Fig. 4), simply because the fitnesses of $W_{1/+}$ and $W_{1/2}$ must then be close to each other if an equilibrium is to be possible at all. Conversely, the estimate becomes inaccurate when $W_{2/+}$ is fit (lower right of Fig. 4). Fig. 5 shows the sampling error in the estimate of the geometric mean fitness $\sqrt{(W_{1/+}\,W_{2/+})}/W_{1/2}$, which is a measure of the average fitness effect of the + chromosome. This is most accurate when fitnesses are equal; then, the sampling variance is $9/2n$.

Sampling drift will also generate errors in the estimates. Suppose that the sampling of N zygotes to found each generation causes a random variance ϵ_{ij} in the genotype frequencies, u_i. Let V_{ij} be the covariance in genotype frequencies produced by the accumulation of these fluctuations, and let u_i^* be the deterministic expression for genotype frequencies after one generation. The matrix equation $V_{ij} = \partial_{uk}\, u_i^*\, \partial_{ul}\, u_j^*\, V_{kl} + \epsilon_{ij}$ can be solved to give the variance in genotype frequencies, which in turn gives the variances of the estimated fitnesses. (Note that sampling zygotes is not quite equivalent to sampling gametes, as in the Wright–Fisher model.) For the symmetrical case ($u_{1/+} = u_{2/+} = u_{+/+} = 1/3$), the variances of both $W_{1/+}/W_{1/2}$ and $W_{2/+}/W_{1/2}$ are $136/9N$, and the covariance between these estimates is $116/9N$. Thus, the cumulative effects of sampling zygotes causes $(136/9N/6/n) \sim 2{\cdot}5\,n/N$ times as much variance in each estimate as the immediate sampling error, and also causes a stronger correlation between the estimates. However, if samples are much smaller than the effective number of zygotes ($n \leqslant N$), random drift will cause negligible error compared with random sampling.

3. Displacement of one balancer by another

If the balancers differ substantially in their effects on fitness, then a stable equilibrium in which all three chromosomes segregate is impossible (Fig. 2); this was the case for the TM1 and TM2 balancers used by

https://doi.org/10.1017/S0016672399004346 Published online by Cambridge University Press

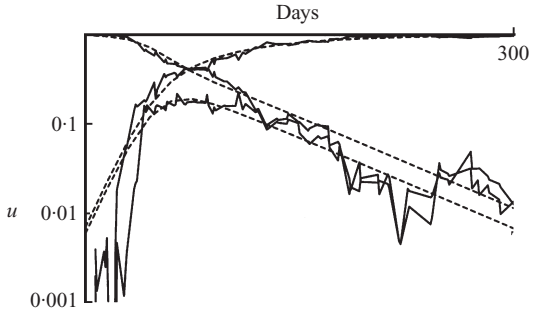

Fig. 6. The proportions, u, of each genotype observed in cage 5A of Fowler *et al*. (1997), plotted against time on a logarithmic scale. The cage initially contained the TM1/+ genotype. Chromosome TM2 was introduced at low frequency, and displaced TM1 to give a cage composed almost entirely of TM2/+ after 300 days. The balancer heterozygote TM1/TM2 was found only during the course of the invasion. The dashed lines show the maximum likelihood estimate: an initial frequency $p_2 = 0{\cdot}0050$, fitnesses $W_{1/+} = 1{\cdot}11\ W_{1/2}$, $W_{2/+} = 2{\cdot}67\ W_{1/2}$ and viabilities in the sample vials of $V_{1/+} = 1{\cdot}68\ W_{1/2}$, $V_{2/+} = 1{\cdot}18\ W_{1/2}$.

Fowler *et al*. (1997). However, relative fitnesses can still be measured using a different design. Suppose that B2/+ is substantially fitter than B1/+, and that B2 chromosomes are introduced at low frequency into a population which is at a stable equilibrium for B1/+ and +/+. Then, B2 will increase exponentially, and will eventually displace B1 to give a population containing only B2/+ and +/+. The standard genotype B1/B2 will be formed transiently, during the replacement of one balancer by the other. This design has the drawback that multiple samples must be taken throughout the invasion of B1. Also, relative fitnesses might change through time, as a result of direct or indirect frequency-dependence. However, there are advantages. First, because there is information on changes through time, it is possible to estimate the relative fitnesses of the four genotypes that are present when +/+ are viable. In the simple equilibrium case, zygote frequencies depend on the proportions of the three chromosomes, and so only two degrees of freedom are available. Secondly during the initial and final stages of the invasion, the rare chromosome is present within two heterozygous genotypes. If +/+ are lethal, the proportions of these must be the same in zygotes; hence, relative viabilities in the sample vials can be estimated without the need for calibration by separate experiments (Fowler *et al*., 1997). Finally, observations on a population which is changing through time allow adult lifespan to be measured (see below).

Fowler *et al*. (1997) estimated relative fitnesses by maximum likelihood. The model was specified using the discrete generation recursion, (1), and was then interpolated to continuous time, assuming some plausible generation time τ. (We consider more realistic age-structured models below.) The likelihood was calculated assuming binomial sampling error, and maximized using the Newton–Raphson algorithm. As an example, consider cage 5A (i.e. the first replicate cage containing wild-type chromosomes #5; Fig. 6). Using the full model, and assuming a generation time of 15 days, the maximum likelihood estimates are

$$W_{1/+} = 1{\cdot}11\ W_{1/2}\ (\text{SE } 0{\cdot}037),$$
$$W_{2/+} = 2{\cdot}67\ W_{1/2}\ (\text{SE } 0{\cdot}052).$$

The standard errors of these estimates were found from the inverse of the information matrix, which is minus the second differential of the log likelihood (Edwards, 1972). These sampling errors are relatively small, reflecting the large numbers of flies that were counted (for this cage, 61047 flies). The two fitness estimates are strongly correlated ($r = 89\,\%$), which increases the error in estimating the mean effect of this wild-type chromosome ($\sqrt{W_{1/+}\,W_{2/+}}/W_{1/2} = 1{\cdot}72$, SE 0·044).

To understand the relation between the observed time course and the estimated fitnesses, it is helpful to use a simple argument based on the initial rate of increase of B2, and the final rate of elimination of B1. For simplicity, suppose that +/+ is lethal. Then, B2 is initially found in equal proportions in B2/+ and in B1/B2; it increases by a factor

$$e^{\lambda_0 \tau} = \frac{(W_{2/+} + W_{1/2})}{W_{1/+}}$$

in each generation. For example, if all heterozygotes have the same fitness, then the rare chromosome doubles in frequency in every generation. Conversely, the final elimination of B1 occurs at a rate

$$e^{\lambda_1 \tau} = \frac{(W_{1/+} + W_{1/2})}{W_{2/+}}$$

Thus, the relative fitnesses can be estimated from the initial and final rates of change. Note, however, that if B1/B2 has low fitness, both λ_0 and λ_1 are given approximately by the ratio between the fitness of B1/+ and B2/+. It is then difficult to disentangle the fitnesses relative to the reference genotype B1/B2 from these two rates.

To compare this simplified method with the full maximum likelihood estimate, consider the example of Fig. 6. The fitnesses estimated from the full model ($W_{1/+} = 1{\cdot}11\ W_{1/2}$, $W_{2/+} = 2{\cdot}67\ W_{1/2}$) correspond to initial and final growth rates of $\lambda_0 = 0{\cdot}080$, $\lambda_1 = -0{\cdot}0157$. Now, suppose we estimate the initial and final rates λ_0, λ_1 simply by using date from the periods up to day 35, and after day 100, and assuming that the observed numbers of the rare chromosome follow a Poisson distribution. This gives $\lambda_0 = 0{\cdot}174$ (SE 0·0094) and $\lambda_1 = -0{\cdot}0149$ (SE 0·00042). These correspond to increase by a factor 13·6, 0·80 respectively, over a 15

https://doi.org/10.1017/S0016672399004346 Published online by Cambridge University Press

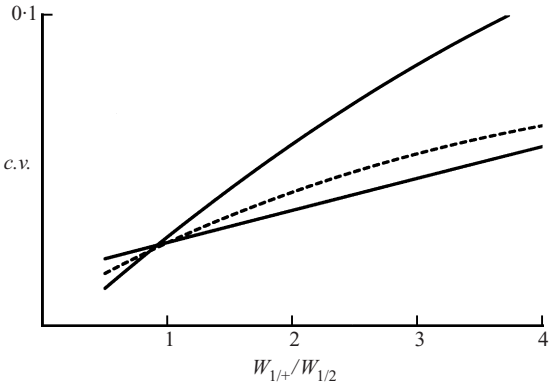

Fig. 7. The coefficient of variation (*c.v.*; i.e. standard error/mean) of fitness estimates made using the invasion method. The upper line shows the *c.v.* of $W_{2/+}/W_{1/2}$, and the lower line, the *c.v.* of $W_{1/+}/W_{1/2}$. The dashed line shows the *c.v.* of estimates of $\sqrt{(W_{1/+}\, W_{2/+})}/W_{1/2}$. Values are for 50 samples of 1000 individuals, evenly spaced over 300 days; in general, $(c.v.)^2$ is inversely proportional to sample size. The generation time is 15 days, the initial frequency of B2 is 0·0005, and $W_{2/+} = 2\, W_{1/+}$. For $W_{1/+}/W_{1/2} < 1$, the population evolves towards a stable polymorphism for all three chromosomes, whilst for $W_{1/+}/W_{1/2} > 1$, B2 replaces B1.

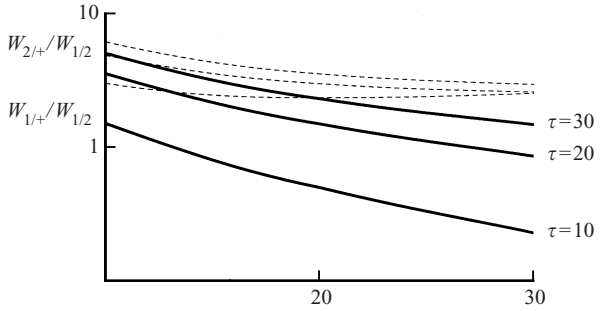

Fig. 8. The continuous lines show the maximum likelihood estimates of $W_{1/+}/W_{1/2}$, plotted on a log scale against the generation time assumed in the fit (in days). The three lines correspond to three datasets, generated from models with generation times $\tau = 10$, 20, 30, with fitnesses $W_{1/+} = 1{\cdot}5\, W_{1/2}$, $W_{2/+} = 3W_{1/2}$, and initial frequency 0·0005 of B2. Dashed lines show estimates of $W_{2/+}/W_{1/2}$, and are in the same order. Note that when the fitted model corresponds to the model that generated the data, the estimated fitnesses coincide with the true ones. Increasing the assumed generation time entails larger fitness differences.

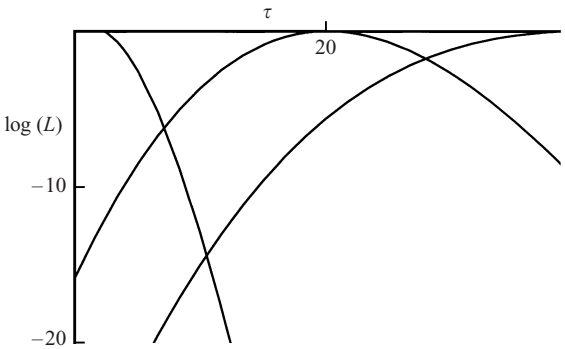

Fig. 9. The log(likelihood) of the models fitted in Fig. 8, plotted against the generation time assumed in the fit (in days). The three lines correspond to three datasets, generated from models with generation times $\tau = 10$, 20, 30, with fitnesses $W_{1/+} = 1{\cdot}5\, W_{1/2}$, $W_{2/+} = 3\, W_{1/2}$, and initial frequency 0·0005 of B2. Fifty samples, each of 1000 individuals, are spread evenly over 300 days, as in Fig. 7. When the fitted model corresponds to the model that generated the data, the fit is perfect, and the log(likelihood) reaches a maximum of zero.

day generation time, and give estimates of $W_{1/+} = 0{\cdot}18\, W_{1/2}$ (SE 0·028), $W_{2/+} = 1{\cdot}48\, W_{1/2}$ (SE 0·037). These two fitness estimates are almost completely correlated with each other ($r = 0{\cdot}966$); the estimate of the geometric mean fitness is $\sqrt{W_{1/+}\, W_{2/+}}/W_{1/2} = 0{\cdot}52$, SE 0·047). The reason why these estimates are so different is that the best estimate for the full model is for a higher initial frequency, and a slower initial rate of increase (left of Fig. 6): though this gives a poorer fit in the early stages, it accounts better for the intermediate period, during which all three chromosomes are common. Fitting the full model by maximum likelihood incorporates additional information from the intermediate period during which all three chromosomes are at high frequency.

Fig. 7 shows the coefficient of variation associated with the fitness estimates based on the full model, and plotted against $W_{1/+}/W_{1/2}$; in this example, $W_{2/+} = 2W_{1/+}$. The coefficient of variation was calculated from information matrix, assuming that the data follow the theoretical expectation. Results are for 50 samples of 1000 individuals (similar to the number used by Fowler *et al.*, 1997), spread evenly over 300 days (here, 20 generations). The coefficient of variation scales inversely with the square root of sample size, which allows values to be calculated for other sample sizes. Sampling errors are small when fitnesses are similar to each other (left of Fig. 7). However, if B1/B2 is very unfit, estimates become inaccurate, since the pattern depends primarily on $W_{1/+}/W_{2/+}$, rather than on their separate values (see above).

If an equilibrium is reached, then the relative fitnesses can be estimated without the need for any assumptions as to generation time. However, the generation time determines the rate at which one chromosome displaces another, and hence influences fitness estimates made from this rate. With a short generation time, smaller fitness differences are needed to account for a given pattern, whilst with a long generation time, larger fitness differences will be estimated (Fig. 8). In principle, the generation time could itself be estimated. Fig. 9 shows the log(likelihood) of fits to three sets of data, each generated so as to follow precisely models with generation time s$\tau = 10$, 20 or 30 days. The likelihood is necessarily maximized at the true value, and decreases if an inappropriate value is assumed.

https://doi.org/10.1017/S0016672399004346 Published online by Cambridge University Press

However, even though a total sample size of 50000 individuals is used in this example, the change in likelihood is modest. In practice, systematic deviations from the idealised discrete model, and other sources of random variation, would make it impossible to estimate generation time in this way. Indeed, the actual patterns estimated using different generation times are almost indistinguishable.

4. Age structure

Thus far, we have fitted a simple model of reproduction in discrete generations. In fact, *Drosophila* population cages are age-structured: at any time, all stages in the life cycle may be present. At equilibrium, it is not obvious how net fitness is composed from each fitness component, or how it corresponds to fitnesses estimated using the discrete model. In the invasion design, there is the further complication that age structure may change through time. Here, we address these issues using a simple model of an age-structured population. Only the case where the wild-type homozygote is lethal will be considered.

First, consider a single heterozygous genotype. The number of adults present at time t is $n[t]$; adults are assumed to have a constant mortality rate d. For simplicity, we assume an equal sex ratio at birth, and equal mortality for males and females. Each of the $n[t]/2$ females produces f eggs per day; of these, half will be homozygous, and so will die at an early stage. In order for population size to approach an equilibrium, fitness must depend on density: we assume that fitness decreases exponentially with density, and that it depends on the numbers present in the early larval stage, though after homozygotes have died. This introduces a term $\mathrm{e}^{-fn[t-T]/(4K)}$, where K is proportional to carrying capacity. The time from birth to eclosion is T days. Thus, the proportion surviving to eclosion at time t is $v\mathrm{e}^{-fn[t-T]/(4K)}$, since $n[t-T]f/4$ surviving eggs are laid at time $t-T$. (It would be more realistic to suppose that the larval viability depends on the numbers of surviving larvae over a range of ages; the crude approximation is made here to avoid introducing another parameter into the model, and to avoid integral equations.) The net rate of change in adult numbers is:

$$\frac{\mathrm{d}n}{\mathrm{d}t} = -\mathrm{d}n[t] + \frac{fv}{4}\mathrm{e}^{-fn[t-T]/(4K)}\,n[t-t]. \tag{7}$$

The equilibrium density of adults is:

$$n_{\mathrm{eqbrm}} = \frac{4K}{\mathrm{f}}\log\left[\frac{vf}{4d}\right]. \tag{8}$$

The equilibrium density decreases with adult mortality (d), and increases with larval viability (v); however, the dependence is only logarithmic. The influence of female fecundity is counter-intuitive: numbers at first increase with f, but then decrease as eggs become too crowded. Fig. 10 shows how the numbers of larvae and adults approach equilibrium, after introduction of a small number of eggs. The oscillations decay as the age structure approaches its equilibrium. Since the discrete-generation analogue of this model of density-dependence gives cyclic or chaotic dynamics for sufficiently high rates of population growth, the same may be true of this age-structured version; we do not investigate this issue here.

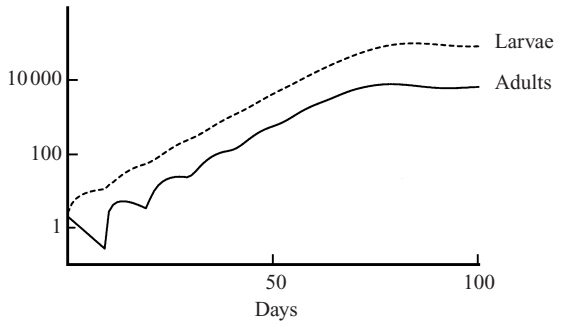

Fig. 10. The numbers of adults and larvae plotted on a log scale against time. Two adults were introduced; adult numbers declined exponentially (continuous line, lower left), while the numbers of the larvae they produced increased (dashed line, lower left). Numbers of adults increased sharply when the first offspring eclosed, and subsequently increased towards equilibrium. Larval viability $v = 0{\cdot}05$, adult death rate $d = 0{\cdot}2$ per day, development time $T = 10$ days, female fecundity 100 eggs per day, $\kappa = 10^5$, giving an equilibrium density $n_{\mathrm{eqbrm}} =$ 7330 adults.

The model extends to include three heterozygous genotypes by assuming that each genotype has its own life history, characterized by larval survival (v); development time (T); density-dependence (κ); adult mortality (d); and female fecundity (f). An additional parameter (m) must be included for each genotype to describe its male fertility. Females are assumed to lay f eggs per day regardless of who they mate with, whilst males are assumed to contribute in proportion to m. The rate of change of the number of B1/B2 adults is:

$$\frac{\mathrm{d}n_{12}}{\mathrm{d}t} = -d_{12}\,n_{12}[t] + \frac{\tilde{v}_{12}}{4}\left(F + \frac{n_{12}^2 f_{12} m_{12} - n_{01}^2 f_{01} m_{01} - n_{02}^2 f_{02} m_{02}}{4M}\right),$$

where

$$\tilde{v}_{12} \equiv v_{12}\exp\left[-\frac{1}{2K_{12}}\left(\frac{3F}{2} - \frac{\Sigma_\beta n_\beta^2 f_\beta m_\beta}{8M}\right)\right], \tag{9}$$

$$F \equiv \Sigma_\beta \frac{n_\beta f_\beta}{2}, \quad M \equiv \Sigma_\beta \frac{n_\beta m_\beta}{2}.$$

The first term in (9) is due to the death of B1/B2 adults, whilst the second is due to newly eclosing adults, surviving with probability $\tilde{v}_{12}$ from eggs laid at $t-T_{12}$; the second term is evaluated at $t-T_{12}$. The

https://doi.org/10.1017/S0016672399004346 Published online by Cambridge University Press

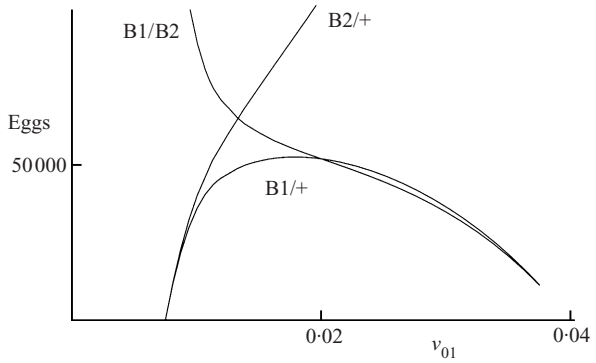

Fig. 11. Equilibrium rate of egg production of the three viable genotypes, plotted against larval viability of B1/+. ν_{01}, in the age-structured model Viability of B1/B2 is fixed at 0·02, and the viability of B2/+ is 1·5 times that of B1/+. Male fertilities are equal, and female fecundities are 100 eggs per day. Carrying capacity $K = 10^5$. Below $\nu_{01} = 0{\cdot}0075$, B1/B2 fixes, whilst above $\nu_{01} = 0{\cdot}039$, B2/+ fixes.

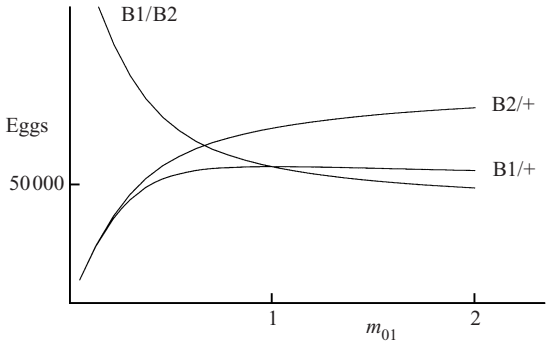

Fig. 12. Equilibrium rate of egg production of eggs of the three viable genotypes, plotted against fertility of male B1/+, m_{01}, in the age-structured model. Fertility of B1/B2 is fixed at 1, and the fertility of B2/+ is 1·5 times that of B1/+. Larval viabilities are 0·02, and female fecundities are 100 eggs per day. Carrying capacity $K = 10^5$.

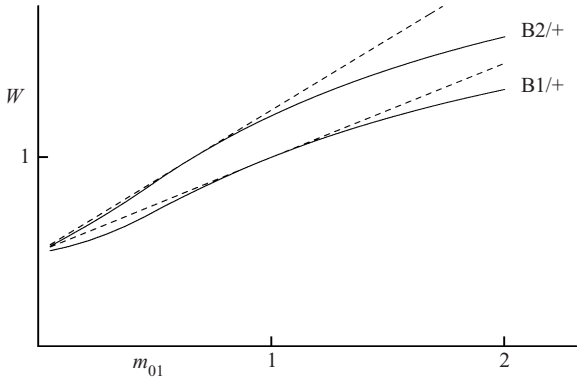

Fig. 13. Continuous curves show fitnesses of B2/+ and B1/+, relative to B1/B2, estimated using the discrete model from egg numbers generated by the age-structured model. These are plotted against the fertility of male B1/+; all parameters are as in Fig. 12. The dashed lines show the naive approximation that fitness is proportional to the average of male fertility and female fecundity.

larval viability $\tilde{v}_{12}$ is reduced from v_{12} by an exponential factor, $\exp[-F^*/K_{12}]$, where F^* is the *effective rate* of production of eggs. This is smaller than the total rate of egg production, F, because homozygous genotypes are assumed to die young, and not to compete with potentially viable larvae. The factor $(F+(\ldots)/(4M))$ is found by summing over the contributions from all possible matings between the three kinds of male and female.

The equilibria can be found numerically from (9). If male fertility and female fecundity are proportional to each other ($m_\alpha \sim f_\alpha$), then the equilibrium takes the same form as for the discrete model, with fitness measure $\tilde{W}_\alpha = \tilde{\nu}_\alpha f_\alpha / d_\alpha$; note that the larval viability $\tilde{\nu}_\alpha$ is density- and frequency-dependent. This measure is just the lifetime reproductive success of each genotype. Fig. 11 shows the equilibrium rate of production of eggs of the three viable genotypes, as the viability of B1/+ and B2/+ increases; the viability of the reference genotype B1/B2 remains fixed, and B2/+ has viability 50% greater than that of B1/+. If the viabilities of B1/+, B2/+ are below a critical threshold, B1/B2 is fixed (left of Fig. 11), while above a threshold, B2/+ fixes (right of Fig. 11). For this example, the numbers of adults show a similar pattern. Net egg production (the sum of the three curves in Fig. 11), and net adult numbers, increase with larval viability, as expected. Estimates of fitness made from these zygote numbers, using the discrete model, would give values proportional to the larval viability. More generally, the discrete model would lead to correct fitness estimates provided that male fertility and female fecundity are proportional to each other.

However, if selection acts differently on males and females, no single genotypic fitness measure suffices to define the zygote frequencies, since these will not be in Hardy–Weinberg proportions: male and female fitnesses must be considered separately. Fig. 12 shows how the rate of egg production varies with male fertility. The net rate of egg production now barely changes with male fertility, since the number of eggs produced by each female is not directly affected by her mate. (There is an indirect effect, however, since if she mates with the same genotype, half her offspring are viable, whereas if she mates with a different genotype, three-quarters are viable.) If B1/B2 is much more fertile than the other genotypes, it approaches fixation (left of Fig. 12), whilst if it is much less fertile, a polymorphism with all three genotypes is approached (right of Fig. 12). For these parameters, there is no threshold, because genotypes with low male fertility still contribute female fecundity. Fig. 13 shows the relative fitnesses that would be estimated using the

https://doi.org/10.1017/S0016672399004346 Published online by Cambridge University Press

Table 1. *Estimates of fitness for the age-structured model, with selection on larval viability* ($\nu_{1/+}$, $\nu_{1/2}$, $\nu_{2/+}$)

$v_{1/+}$	$v_{1/2}$	$v_{2/+}$	$W_{1/+}$	$W_{1/2}$	$W_{2/+}$	log(L)	λ_0	λ_1	$\hat{W}_{1/+}$	$\hat{W}_{1/2}$	$\hat{W}_{2/+}$	log($\hat{L}$)
0·01	0·01	0·02	1·21	1	2·11	−51·34	0·065	0	1·21	1	2·21	−259·42
0·02	0·01	0·04	2·36	1	4·11	−48·98	0·053	−0·015	2·29	1	4·10	−147·11
0·03	0·01	0·06	3·44	1	5·97	−46·86	0·049	−0·020	3·23	1	5·72	−156·07
0·04	0·01	0·08	4·53	1	7·87	−54·81	0·047	−0·023	4·04	1	7·15	−157·42

Data were generated by taking 51 samples of 1000 viable zygotes, evenly spread over 300 days; the initial frequency of B2 was $p_0 = 5 \times 10^{-4}$. $W_{1/+}$, $W_{1/2}$, $W_{2/+}$ give the fitnesses estimated by maximum likelihood from the discrete model, assuming a generation time $\tau = 15$ days; log(L) gives the log(likelihood) for this fit. λ_0 is the initial rate of increase of B2, and λ_1 the final rate of decline of B1 (from Eq. 10). $\hat{W}_{1/+}$, $\hat{W}_{1/2}$, $\hat{W}_{2/+}$ are the fitnesses estimated from λ_0, λ_1, by setting $e^{\lambda_0 \tau} = (\hat{W}_{1/2} + \hat{W}_{2/+})/\hat{W}_{2/+}$, $e^{\lambda_1 \tau} = (\hat{W}_{1/2} + \hat{W}_{1/+})/\hat{W}_{2/+}$. Adult mortality is $d = 0{\cdot}1$ per day, development time 10 days, carrying capacity $K = 10^5$, and female fecundity $f = 100$ eggs per day; male fertilities are equal ($m = 1$).

discrete model, based on egg numbers generated from the age-structured model of Fig. 12. There is no simple relation between the male fertility and the estimated fitness. The naive approximation that fitness is the arithmetic average of male fertility and female fecundity (dashed lines) fits well when B1/B2 is the fittest genotype, but overestimates fitness differences when B1/B2 is unfit. If selection on males and females differs, then the evolution of the population depends on both male and female contributions, and cannot be described by a single composite measure which can be estimated from data pooled across sexes (Nagylaki, 1990).

Next, consider the invasion method, in which genotype B2/+ displaces B1/+. The criteria for invasion of the population differ from those for equilibrium. In the discrete model, the rate of increase of B2, λ_0, is given by $e^{\lambda_0 \tau} = (W_{2/+} + W_{1/2})/W_{1/+}$, where τ is the generation time. The corresponding expression for λ_0 in the age-structured model is:

$$\frac{\tilde{v}_{01}}{4\,d_{01}} = \frac{\tilde{v}_{12}\, e^{-\lambda_0 T_{12}}}{4(\lambda_0 + d_{12})}\left(\frac{f_{12}}{2f_{01}} + \frac{m_{12}}{2m_{01}}\right) \times \frac{\tilde{v}_{02}\, e^{-\lambda_0 T_{02}}}{4(\lambda_0 + d_{02})}\left(\frac{f_{02}}{2f_{01}} + \frac{m_{02}}{2m_{01}}\right), \quad (10)$$

where

$$\tilde{v}_{12} \equiv v_{12} \exp\left[-\frac{n_{01} f_{01}}{4K_{12}}\right], \quad \tilde{v}_{02} \equiv v_{02} \exp\left[-\frac{n_{01} f_{01}}{4K_{02}}\right].$$

If development times of the increasing genotypes are equal ($T = T_{12} = T_{02}$), the fitness measures correspond if we set:

$$\frac{\hat{W}_{2/+}}{\hat{W}_{1/+}} \equiv \frac{\tilde{v}_{02}/\tilde{v}_{01}}{(\lambda_0 + d_{02})/d_{01}}\left(\frac{f_{02}}{2f_{01}} + \frac{m_{02}}{2m_{01}}\right),$$

$$\frac{\hat{W}_{1/2}}{\hat{W}_{1/+}} \equiv \frac{\tilde{v}_{12}/\tilde{v}_{01}}{(\lambda_0 + d_{12})/d_{01}}\left(\frac{f_{12}}{2f_{01}} + \frac{m_{12}}{2m_{01}}\right). \quad (11)$$

Then, the rate of increase is given by $e^{\lambda_0 T} = (\hat{W}_{2/+} + \hat{W}_{1/2})/\hat{W}_{1/+}$. This differs in two ways from the discrete model. First, the formula involves the time from egg to eclosion, T, rather than the generation time (here, $\tau \sim T + 1/d$). Secondly, the death rates of the increasing genotypes are augmented by the rate of increase of the population ($\lambda_0 + d_{02}$, etc.), because recent births contribute more to an expanding population. If male fertilities are proportional to female fecundities, the equilibrium (if it exists) depends on $\tilde{W}_\alpha \equiv \tilde{\nu}_\alpha f_\alpha / d_\alpha$, while the rate of increase depends on $\hat{W}_\alpha = \tilde{W}_\alpha d_\alpha/(\lambda_0 + d_0)$. Fitness estimates based on the rate of invasion will differ from those based on equilibrium, both because the fitness components actually differ at different densities and genotype frequencies (for example, $\tilde{\nu}$ is density-dependent), and because of the demographic effect of population increase.

In order to compare the discrete and age-structured models we can ask, first, how fitness estimates based on fitting the discrete model relate to the various parameters of the age-structured model, and, second, whether the time course of genotype frequency is distinguishable under the two models. We construct datasets using the age-structured model, by taking 51 samples of 1000 viable zygotes, spread evenly over 300 days. The default parameters are adult mortality $d = 0{\cdot}1$ per day, larval liability $\nu = 0{\cdot}01$, development time 10 days, carrying capacity $K = 10^5$, and female fecundity 100 eggs per day. We then fit the discrete model to these data, assuming (arbitrarily) a generation time of $\tau = 15$ days. Table 1 shows the estimates where larval viability varies between genotypes, with $\nu_{2/+} = 2\,\nu_{1/+}$, and an initial frequency $p_0 = 5 \times 10^{-4}$ of the B2 chromosome. (The initial frequency p_0 was taken as known, to speed calculation). The fitnesses of B1/+ and B2/+ estimated from the discrete model are somewhat greater, relative to B1/B2, then the ratios of viabilities. The fit of the discrete model is indistinguishable from the age-structured model which generated the data: the log(likelihood) of the best fit is close to −49, the value expected from an asymptotic $-\frac{1}{2}\chi^2$ distribution with $2(51-2) = 98$ df.

https://doi.org/10.1017/S0016672399004346 Published online by Cambridge University Press

Table 2. *Estimates of fitness for the age-structured model, with selection on male fertility* ($m_{1/+}$,$m_{1/2}$,$m_{2/+}$)

$m_{1/+}$	$m_{1/2}$	$m_{2/+}$	$W_{1/+}$	$W_{1/2}$	$W_{2/+}$	log(L)	λ_0	λ_1	$W_{1/+}$	$W_{1/2}$	$W_{2/+}$	log(L)	$\tilde{W}_{1/+}$	$\tilde{W}_{1/2}$	$\tilde{W}_{2/+}$	log(L)
1	1	2	1·07	1	1·40	−58·51	0·053	0·022	1·14	1	1·53	−65·78	1	1	1·37	−175·21
2	1	4	1·39	1	1·80	−53·61	0·047	0·017	1·41	1	1·87	−58·94	1·33	1	1·75	−88·34
3	1	6	1·57	1	2·03	−53·84	0·044	0·016	1·55	1	2·07	−55·86	1·51	1	1·94	−59·32
4	1	8	1·66	1	2·15	−41·77	0·043	0·015	1·62	1	2·10	−43·74	1·62	1	2·05	−62·87

Other parameters are as for Table 1, except that larval viability is set to $\nu = 0{\cdot}01$. The additional four columns show the fitnesses estimated from the final equilibrium, $\{\tilde{W}_{1/+}, \tilde{W}_{1/2}, \tilde{W}_{2/+}\}$.

The relative fitnesses estimated from the discrete model are close to those required to give the observed initial rates of increase and decrease, calculated from (10) (last four columns of Table 1). The log(likelihood) is substantially worse, reflecting the large sample size in the simulation. This shows that the discrete model, described by two relative fitnesses, gives a good summary of the full age-structured model, which involves many more parameters. Table 2 shows fits to data generated from a population subject to selection on male fertility rather than larval viability. Since selection now acts only on one sex, a stable equilibrium is possible; fitnesses estimated from this equilibrium are shown in the last four columns. These estimates are quite close to the maximum likelihood estimates, and to those estimated from the rates of invasion of B1 and B2. The pattern is again similar to that seen for selection on larval viability (Table 1). Table 3 shows fits to data sampled from a population in which selection acts solely on carrying capacity. This density-dependent selection can lead to initial and final rates which could not be realized in the discrete model: in the examples in Table 1, population size increases as B2/+ replaces B1/+ (Fig. 14). Thus, the final rate of decrease of B1 can be faster, relative to the initial rate of increase, than could be produced even if B1/B2 were completely inviable.

The relative fitnesses differ from the relative viabilities primarily because the population of B2 chromosomes is initially increasing ($\lambda_0 > 0$), and the B1 chromosomes are eventually eliminated ($\lambda_1 < 0$); inspection of the denominators in (10) shows that this introduces factors of $(1+\lambda_0/d), 1/(1+\lambda_1/d)$ into the relative fitnesses of B2/+ and B1/+. The relative fitnesses are also scaled by the arbitrary assumption of a generation time $\tau = 15$ days made when fitting the discrete model. There is no simple fitness measure which precisely describes the effects of the various fitness components on genotype frequencies throughout the invasion. The conventional approach of taking the intrinsic rate of increase of each chromosome (Charlesworth, 1980) applies at the beginning and end of the invasion, when the age structure approaches an asymptotic state. However, this measure of chromosomal fitness involves a combination of the life histories of those genotypes in which it is found, and is not a simple composition of any diploid fitness measures: (11) depends on the rate of increase, λ_0. When all three genotypes are common, and the age structure is changing rapidly, matters are still more complicated. These problems reflect fundamental difficulties in defining fitness in the presence of age structure and density-dependence. Nevertheless, our numerical examples show that the discrete model gives a good overall fit to data generated with age structure, provided that the initial and final rates of increase are in fact feasible under the discrete model.

https://doi.org/10.1017/S0016672399004346 Published online by Cambridge University Press

Table 3. *Estimates of fitness for the age-structured model, with selection on carrying capacity* ($K_{1/+}$,$K_{1/2}$,$K_{2/+}$)

$K_{1/+}$	$K_{1/2}$	$K_{2/+}$	$W_{1/+}$	$W_{1/2}$	$W_{2/+}$	log(L)	λ_0	λ_1	$W_{1/+}$	$W_{1/2}$	$W_{2/+}$	log(L)
10^5	10^5	2×10^5	1·88	1	3·11	−75·94	0·055	−0·011	1·98	1	3·53	−880·21
2×10^5	10^5	4×10^5	9·05	1	14·77	−66·01	0·039	−0·036	40·46	1	71·48	−1195·55
3×10^5	10^5	6×10^5	1	0	1·57	−68·27	0·031	−0·041				
4×10^5	10^5	8×10^5	1	0	1·51	−112·88	0·028	−0·042				

Other parameters are as for Table 1, except that larval viability is set to $v = 0{\cdot}01$. For $K_{1/+} > 2\times10^5$, no set of relative fitnesses can match the initial and final rates, λ_0,λ_1; the best estimate is for $W_{1/+}$,$W_{2/+}$ to be much larger than $W_{1/2}$.

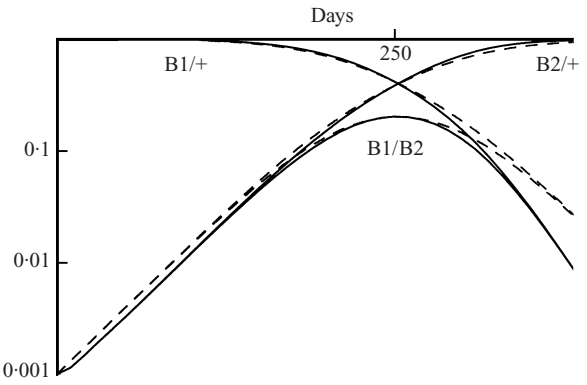

Fig. 14. The frequencies of the three genotypes, for selection on carrying capacity: $K_{1/+} = 4\times10^5$, $K_{1/2} = 10^5$, $K_{2/+} = 8\times10^5$. Larval viability is $v = 0{\cdot}01$; time to eclosion 10 days; adult mortality is $d = 0{\cdot}1$ per day; male fertilities are equal; and female fecundity is 100 eggs per day. Initial frequency of B2 is $p_0 = 0{\cdot}0005$. The continuous lines show the actual time course, and the dashed lines, the best fit under the discrete model: $W_{1/+} = 1$, $W_{1/2} = 0$, $W_{2/+} = 1{\cdot}51$. This figure corresponds to the last row in Table 3.

5. Discussion

(i) *Summary of the methods*

The effects of wild-type chromosomes on heterozygous fitness can be estimated by competing them against two different balancer chromosomes in replicated population cages. If the fitnesses of the three heterozygous genotypes are not too dissimilar, then a polymorphic equilibrium is reached; the equilibrium frequencies then give straightforward estimates of the fitnesses of the two kinds of wild-type heterozygotes. These are measured relative to the B1/B2 heterozygote, which acts as a standard reference genotype. Fowler *et al*. (1997) found that the two balancers had sufficiently different effects that no equilibrium could be reached; fitnesses can then be estimated from the time course of an invasion in which one wild-type heterozygote displaces the other. We now discuss briefly some strengths and weaknesses of this approach, and ways in which it might be elaborated to investigate fitness components. A *Mathematica* notebook containing the algorithms used to implement these methods can be found at http://helios.bto.ed.ac.uk/evolgen/index.html.

(ii) *Additivity, epistasis and frequency-dependence*

Our method provides separate estimates of the fitnesses of the two kinds of wild-type heterozygote. In principle, these can be separated into additive effects of the + chromosome, and a non-additive interaction with the balancer genotype. However, this does not relate in a simple way to the additive and dominance effects which would be found within a wild-type population. Moreover, because the two fitness estimates are strongly correlated, it is in practice hard to separate the two components (see Fowler *et al*., 1997).

Intra-chromosomal epistasis present in the wild-type chromosomes at extraction (Charlesworth & Barton, 1996) will contribute to the fitness estimate obtained, simply because the effect of the whole chromosome is being measured. The method could be extended to measure the magnitude of such epistasis, by observing the total fitnesses of recombinant chromosomes. The fitnesses of the wild-type chromosomes are assessed in a standard genetic background, raising the possibility of inter-chromosomal epistasis as a contributor to the measured fitness effects. However, the genetic background is the random-bred one in the populations from which the chromosomes were extracted, and interactions with it would therefore contribute to fitness in the base chromosome.

A potential weakness of the model is density- and frequency-dependence. Some genotypes are present at high frequencies at equilibrium and at various stages in the invasion, and at different frequencies as the invasion proceeds. Frequency dependence in fitness could therefore affect the results. However, all frequencies are represented in all invasions, so that the method does sample across the range of possibilities. Moreover, dependence of fitness on the environmental conditions make problematic any attempt to summarize the effects of a genotype in a single fitness measure.

https://doi.org/10.1017/S0016672399004346 Published online by Cambridge University Press

(iii) *Potential problems with long-term cage experiments*

Mutation accumulation, recombination with the balancers and contamination of the experimental populations are all potential problems with long-term experiments of this kind. Mutation accumulation would occur independently on replicate chromosomes, and would lead to inflated variance between replicate invasions. Large size of the experimental populations would help protect against it and its importance would be much less in populations where the wild-type chromosomes are viable and fertile. The experiments should also be begun as soon as the chromosomes are extracted. Recombination with balancers can be detected by the recessive markers present. It would also tend to deflate variance in the fitness measures obtained, by increasing variability between replicates. Contamination would be immediately apparent, since it would lead to invasion of the experimental cages by wild-type flies.

Genetic drift within each experimental population will affect the frequency of the standard B1/B2 genotype against which the fitnesses of the other genotypes are measured. This will tend to generate spurious correlations in the estimated fitnesses of the two heterozygous frequencies of the wild-type chromosome with the balancer. However, the calculations above show that sampling drift is unlikely to be a serious problem; indeed, Fowler *et al.* (1997) found that the main variation was due to fluctuations in actual fitness peculiar to each replicate cage, which may be due to interactions between genotype and local environment. It is crucial that replicate experiments are run with each chromosome: the main strength of our method is that any differences common to replicate cages carrying the same wild-type chromosome must reflect real effects of that chromosome.

(iv) *The optimal sampling scheme*

It would be interesting to find the optimal intensity of sampling, for a given total sample size. Large samples are required at the beginning and end of the invasion, to give accurate estimates of the frequency of the rare genotypes, and hence of their asymptotic rates of increase and decrease. However, finding the sampling schedule which would minimize sampling error poses a non-trivial numerical problem, and would require sampling to depend on the observed course of the invasion. Fowler *et al.* (1997) showed that the bulk of the errors were due to fluctuations in allele frequency around the fitted model, rather than to sampling. If this extra source of variation depends in the same way on chromosome frequency as does multinomial sampling, then the theory presented above still applies. If it is the main source of variation, accuracy can best be improved by increasing the number of replicate cages, rather than by increasing the sizes of the samples taken from each.

(v) *Fitness with overlapping generations*

We make our estimates by fitting a discrete-generation model, in which fitnesses appear as single parameters. In fact, *Drosophila* cage populations are age-structured, and their evolution is fully described by a set of fitness components which must to some extent depend on population density. If a polymorphic equilibrium is reached, we show that estimates derived from the discrete-generation model bear a simple relation with the fitness components at the equilibrium density. In the invasion experiment, a simple relation between our fitness measure and the fitness components is found only under special circumstances (for example, equal selection on the two sexes), and only during periods of exponential change, in which the age structure has reached a steady state. Moreover, our estimates are conditional on an assumed value for the generation time, which cannot be estimated from genotype frequencies alone. These are inevitable difficulties with defining fitness measures for populations in which the age structure is changing, rather than problems with our particular technique.

(vi) *Measuring fitness components*

It is especially important to separate the various components of fitness for an age-structured population, since no single measure can entirely summarize the action of selection. Fitness components must be measured under the same conditions as net fitness; they can most readily be assayed during the periods of exponential increase and decrease, during which the age structure approaches a steady state.

Male mating success and male fertilization ability can be measured at any time, by comparing genotype frequencies among unmated adults in the cage, mating pairs taken from the cage, parents inferred from the offspring of single mated females, and eggs (see above); all these can be sampled at the same time. In inferring parentage, we assume that only the most recent mate provides sperm. A difficulty is that if the wild-type homozygote is inviable, then any cross between different heterozygotes gives equal proportions of the three heterozygous genotypes in the offspring. However, for those chromosomes (the majority) which are homozygous viable, the father can be inferred for 11 of the 12 possible crosses (a B1/B2 female would give all three heterozygotes whether mated with either B1/+ or B2/+ males). Furthermore, even where the wild-type homozygote is inviable or sterile, matings between the same two heterozygous genotypes can be detected. Provided

https://doi.org/10.1017/S0016672399004346 Published online by Cambridge University Press

that male mating success does not differ with different genotypes of females, this would allow the male fertility component to be measured.

Female fertility would be best assessed by examination of the ovaries of females taken directly from the experimental cages. The ovarian potential of a female can be measured by assessing the presence of eggs and oocytes at different stages of development (Bouletreau, 1978). Larval and adult death rates must be measured through time, while the genotype frequencies are changing. This can be done by sampling pupae and adults during the periods of exponential change. (Note that Fowler *et al.* (1997) found that genotype frequencies did change exponentially for substantial periods, as predicted by theory.) Frequencies among eggs, pupae and adults must change exponentially at the same rate, and the ratio between them at any one time gives estimates of relative larval viability and adult longevity: if a rare genotype β increases at a rate λ, and has death rate β^*, then the ratio between its frequency in eclosing pupae and in adults $\beta/(\lambda+\beta)$, where β is the death rate of the common genotype. These measures can be made under crowded cage conditions.

Viability at low density would be measured from the segregation tests used to correct the egg samples (see above), and will show whether viability differences are exaggerated under harsh conditions (Kondrashov & Houle, 1994). In the invasion method, relative viabilities can also be inferred directly from samples taken while one balancer is rare: then, the two rare heterozygous genotypes must segregate at equal frequencies in zygotes, and so any distortion in flies eclosing from sample vials must be due to differential viability. Using this method, Fowler *et al.* (1997) found significant variation in viability, but no correlation with overall fitness estimates.

(vii) *Relation with fitness in nature and in wild-type cages*

Our estimates are intended as measures of the variation in fitness in a wild-type chromosome that has adapted to cage conditions: it is hard to know how this might relate to fitness variation in a natural population. Our estimates will differ to some extent from those for wild-type cages, for two reasons. First, the balancer chromosomes carry markers that have large fitness effects: there may be interactions with these. In principle, one could construct balancers with markers of imperceptible effects – for example, molecular markers. However, scoring very large numbers of flies would then become difficult. Even with such benign markers, approximately one-third of larvae (those carrying homozygous balancers) will die at an early stage, which will disturb the demography in the cage. Nevertheless, we regard our method as a promising approach to measuring net fitness effects of heterozygous chromosomes.

This work was supported by grant GR31/G09821 from the BBSRC, and by fellowships from the Darwin Trust of Edinburgh and the NERC. We are grateful to the referees for their helpful comments on the manuscript.

References

Barrowclough, G. F. & Rockwell, R. F. (1993). Variance of lifetime reproductive success: estimation based on demographic data. *American Naturalist* **141**, 281–295.

Barton, N. H. (1995). A general model for the evolution of recombination. *Genetical Research* **65**, 123–144.

Barton, N. H. & Turelli, M. (1989). Evolutionary quantitative genetics: how little do we know? *Annual Review of Genetics* **23**, 337–370.

Bertram, D. F. & Strathman, R. R. (1998). Effects of maternal and larval nutrition on growth and form of planktotrophic larvae. *Ecology* **79**, 315–327.

Bouletreau, J. (1978). Ovarian activity and reproductive potential in a natural population of *Drosophila melanogaster*. *Oecologia* **35**, 319–342.

Bryant, E. H. & Meffert, L. M. (1998). Quantitative genetic estimates of morphometric variation in wild-caught and laboratory-reared houseflies. *Evolution* **52**, 626–630.

Burt, A. (1995). The evolution of fitness. *Evolution* **49**, 1–8.

Charlesworth, B. (1980). *Evolution in Age-Structured Populations*. Cambridge: Cambridge University Press.

Charlesworth, B. (1987). The heritability of fitness. In *Sexual Selection: Testing the Alternatives* (ed. J. W. Bradbury & M. B. Anderson), pp. 21–40. New York: Wiley.

Charlesworth, B. (1990). Optimization models, quantitative genetics, and mutation. *Evolution* **44**, 520–538.

Charlesworth, B. (1998). The effect of synergistic epistasis on the inbreeding load. *Genetical Research* **71**, 85–89.

Clutton-Brock, T. H. (1988). Reproductive success: studies of individual variation in contrasting breeding systems. In *Reproductive Success* (ed. T. H. Clutton-Brock), pp. 472–485. Chicago: University of Chicago Press.

Coham, F. M., Hoffman, A. A. & Gayley, T. W. (1989). A test of the role of epistasis in divergence under uniform selection. *Evolution* **43**, 766–774.

Cole, I. C. (1954). The population consequences of life history phenomena. *Quarterly Review of Biology* **29**, 103–137.

Coyne, J. A., Barton, N. H. & Turelli, M. (1997). Perspective: a critique of Sewall Wright's shifting balance theory of evolution. *Evolution* **51**, 643–671.

Crow, J. F. (1958). Some possibilities for measuring selection intensities in man. *Human Biology* **30**, 1–13.

Curtsinger, J. W. (1990). Frequency-dependent selection in *Drosophila*: estimation of net fitness in pseudohaploid populations. *Evolution* **44**, 857–869.

De Visser, J., Arjan, G. M. & Hoekstra, R. F. (1998). Synergistic epistasis between loci affecting fitness: evidence in plants and fungi. *Genetical Research* **71**, 39–49.

Edwards, A. W. F. (1972). *Likelihood*. Cambridge: Cambridge University Press.

Elena, S. F. & Lenski, R. E. (1997). Test of synergistic interactions among deleterious mutations in bacteria. *Nature* **390**, 395–398.

Ellegren, H., Gustaffson, L. & Sheldon, B. C. (1996). Sex-ratio adjustment in relation to paternal attractiveness in a wild bird population. *Proceedings of the National Academy of Sciences of the USA* **21**, 11723–11728.

Endler, J. A. (1977). *Geographic Variation and Speciation in Clines*. Princeton: Princeton University Press.
Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Oxford: Oxford University Press.
Fowler, K., Semple, C., Barton, N. H. & Partridge, L. (1997). Genetic variation for total fitness in *Drosophila melanogaster*. *Proceedings of the Royal Society of London, Series B* **264**, 191–199.
Fox, C. W. & Savalli, U. M. (1998). Inheritance of environmental variation in body size: superparasitism of seeds affects progeny and grandprogeny body size via a nongenetic maternal effect. *Evolution* **52**, 172–182.
Gustaffson, L. & Sutherland, W. J. (1988). The cost of reproduction in the collared flycatcher *Fidedula albicollis*. *Nature* **335**, 813–815.
Houle, D. (1991). Genetic covariance of fitness correlates: what genetic correlations are made of and why it matters. *Evolution* **45**, 630–648.
Houle, D. (1992). Comparing evolvability and variability of quantitative traits. *Genetics* **130**, 195–204.
Houle, D., Hughes, K. A., Hoffmaster, D. K., Ihara, J., Assimakopoulos, S., Canada, D. & Charlesworth, B. (1994). The effects of spontaneous mutation on quantitative traits. 1. Variance and covariance of life history traits. *Genetics* **138**, 773–785.
Houle, D., Morikawa, B. & Lynch, M. (1996). Comparing mutational variabilities. *Genetics* **143**, 1467–1483.
Keightley, P. D. (1996). Nature of deleterious mutations load in *Drosophila*. *Genetics* **144**, 1993–1999.
Keightley, P. D. & Caballero, A. (1997). Genomic mutation rate for fitness in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the USA* **94**, 3823–3827.
Kirkpatrick, M. & Barton, N. H. (1997). The strength of indirect selection on female mating preferences. *Proceedings of the National Academy of Sciences of the USA* **94**, 1282–1286.
Kirkpatrick, M. & Lande, R. (1989). The evolution of maternal characters. *Evolution* **43**, 485–503.
Kirkpatrick, M. & Ryan, M. J. (1991). The evolution of mating preferences and the paradox of the lek. *Nature* **350**, 33–38.
Kondrashov, A. S. (1993). Classification of hypotheses on the advantage of amphimixis. *Journal of Heredity* **84**, 372–387.
Kondrashov, A. S. & Houle, D. (1994). Genotype–environment interactions and the estimation of the genomic mutation-rate in *Drosophila melanogaster*. *Proceedings of the Royal Society of London, Series B* **258**, 221–227.
Latter, B. D. H. & Mulley, J. C. (1995). Gentle adaptation to captivity and inbreeding depression in small laboratory populations of *Drosophila melanogaster*. *Genetics* **139**, 255–266.
Lenski, R. E., Mongold, J. A., Sniegowski, P. D., Travisano, M., Vasi, F., Gerrish, P. J. & Schmidt, T. M. (1998). Evolution of competitive fitness in experimental populations of *E. coli*: What makes one genotype a better competitor than another? *Antonie van Leeuwenhoek International Journal of General and Molecular Microbiology* **73**, 35–47.
Lewontin, R. C. (1965). Selection for colonizing ability. In *The Genetics of Colonizing Species* (ed. H. G. Baker & G. L. Stebbins), pp. 77–94. New York: Academic Press.
Lynch, M. (1985). Spontaneous mutation for life history characters in an obligate parthenogen. *Evolution* **39**, 804–818.
Lynch, M. & Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Sunderland, MA: Sinauer Associates.
Mackay, T. F. C. (1986). A quantitative genetic analysis of fitness and its components in *Drosophila melanogaster*. *Genetical Research* **47**, 59–70.
Maynard Smith, J. (1978). *The Evolution of Sex*. Cambridge: Cambridge University Press.
Mongold, J. A., Bennett, A. F. & Lenski, R. E. (1996). Evolutionary adaptation to temperature. 4. Adaptation of *Escherichia coli* at a niche boundary. *Evolution* **50**, 35–43.
Mukai, T., Chigusa, S. I., Mettler, L. E. & Crow, J. F. (1972). Mutation rate and dominance of genes affecting viability in *Drosophila melanogaster*. *Genetics* **72**, 333–355.
Nagylaki, T. (1992). *Introduction to Theoretical Population Genetics*. Berlin: Springer.
Newton, I. (1989). *Lifetime Reproduction in Birds*. London: Academic Press.
Norris, K. (1993). Heritable variation in a plumage indicator of viability in great tits. *Parus major*. *Nature* **362**, 537–539.
Partridge, L. (1983). Non-random mating and offspring fitness. In *Mate Choice* (ed. P. Bateson), pp. 227–255. Cambridge: Cambridge University Press.
Partridge, I. & Barton, N. H. (1993). Optimality, mutation and the evolution of ageing. *Nature* **362**, 305–311.
Partridge, L. & Harvey, P. H. (1988). The ecological context of life history evolution. *Science* **241**, 1449–1455.
Partridge, L., Mackay, T. F. C. & Aitken, S. (1985). Male mating success and fertility in *Drosophila melanogaster*. *Genetical Research* **46**, 279–285.
Petrie, M. (1994). Improved growth and survival of offspring of peacocks with more elaborate tails. *Nature* **371**, 598–599.
Price, T. & Schluter, D. (1991). On the low heritability of life-history traits. *Evolution* **45**, 853–861.
Prout, T. & McChesney, F. (1985). Competition among immatures affects their adult fertility: population dynamics. *American Naturalist* **126**, 521–558.
Rice, W. R. (1998). Male fitness increases when females are eliminated from gene pool: implications of the Y chromosome. *Proceedings of the National Academy of Sciences of the USA* **95**, 6217–6221.
Rose, M. R. (1984). Laboratory evolution of postponed senescence in *Drosophila melanogaster*. *Evolution* **38**, 1004–1010.
Rowe, L. & Houle, D. (1996). The lek paradox and the capture of genetic variance by condition dependent traits. *Proceedings of the Royal Society of London, Series B* **263**, 1415–1421.
Shabalina, S. A., Yamplosky, L. Y. & Kondrashov, A. S. (1997). Rapid decline of fitness in panmictic populations of *Drosophila melanogaster* maintained under relaxed natural selection. *Proceedings of the National Academy of Sciences of the USA* **94**, 13034–13039.
Sheldon, B. C., Rasanen, K. & Dias, P. C. (1997*a*). Certainty of paternity and paternal effort in the collared flycatcher. *Behavioural Ecology* **8**, 421–428.
Sheldon, B. C., Merila, J., Qvarnstrom, A., Gustafsson, L. & Ellegren, H. (1997*b*). Paternal genetic contribution to offspring condition predicted by size of male secondary sexual character. *Proceedings of the Royal Society of London, Series B* **264**, 297–302.
Sved, J. A. (1971). An estimate of heterosis in *Drosophila melanogaster*. *Genetical Research* **18**, 97–105.
Sved, J. A. (1975). Fitness and third chromosome homozygotes in *Drosophila melanogaster*. *Genetical Research* **25**, 197–200.

Sved, J. A. & Ayala, F. J. (1970). A population cage test for heterosis in *Drosophila pseudoobscura*. *Genetics* **66**, 97–118.

Vassilieva, L. L. & Lynch, M. (1999). The rate of spontaneous mutation for life-history traits in *Caenorhabditis elegans*. *Genetics* **151**, 119–129.

Whitlock, M. C., Phillips, P. C., Moore, F. B. G. & Tonsor, S. J. (1995). Multiple fitness peaks and epistasis. *Annual Review of Ecology and Systematics* **26**, 601–629.

Williams, G. C. (1957). Pleiotropy, natural selection, and the evolution of senescence. *Evolution* **11**, 398–411.

Williams, G. C. (1966). Natural selection, the costs of reproduction, and a refinement of Lack's principle. *American Naturalist* **100**, 687–690.

Williams, G. C. (1975). *Sex and Evolution*. Princeton: Princeton University Press.

Wilton, A. N. & Sved, J. A. (1979). X-chromosomal heterosis in *Drosophila melanogaster*. *Genetical Research* **34**, 303–315.

Zwaan, B. J., Bijlsma, R. & Hoekstra, R. F. (1995). Direct selection of lifespan in *Drosophila melanogaster*. *Evolution* **49**, 649–659.