Genetic variation for total fitness in *Drosophila melanogaster*

KEVIN FOWLER*[1], COLIN SEMPLE‡, NICHOLAS H. BARTON AND LINDA PARTRIDGE*

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

SUMMARY

We measured the heterozygous effects on net fitness of a sample of 12 wild-type third chromosomes in *D. melanogaster*. Effects on fitness were assessed by competing the wild-type chromosomes against balancer chromosomes, to prevent the production of recombinants. The measurements were carried out in the population cage environment in which the life history had been evolving, in an undisturbed population with overlapping generations, and replicated measurements were made on each chromosome to control for confounding effects such as mutation accumulation. We found significant variation among the wild type chromosomes in their additive genetic effect on net fitness. The system provides an opportunity to obtain an accurate estimate of the distribution of heterozygous effects on net fitness, the contribution of different fitness components including male mating success, and the role of intra-chromosomal epistasis in fitness variation.

1. INTRODUCTION

The extent of heritable variation in net fitness determines the response to natural selection, and is crucial for theories of the evolution of sex that rely on short-term benefits (Williams 1975; Maynard Smith 1978). In addition, 'good genes' models of sexual selection require that total fitness should be heritable, and that both male mating success and another component of fitness, such as viability, should contribute to this heritability (Partridge 1983; Charlesworth 1987; Kirkpatrick & Ryan 1991; Burt 1995). There are some theoretical objections to the idea of continuing fitness heritability (Williams 1975; Maynard Smith 1978) and no estimates of the genetic variance of net fitness are available; earlier studies have examined only partial measures of fitness, homozygous effects of chromosomes, or have used an environment other than that in which the life history evolved (Bundgaard & Christiansen 1972; Partridge 1980; Brittnacher 1981; Haymer & Hartl 1981, 1982; Curtsinger 1990; Norris 1993; Petrie 1994). Partial measures of fitness are a problem because different components of fitness are often negatively genetically correlated (Rose 1984*a*; Partridge & Fowler 1992, 1993), while an increase in homozygosity will reveal deleterious effects of recessive alleles that would rarely be expressed in nature (Rose 1984*b*), and a novel environment can induce directional selection (Service & Rose 1985). An adequate approach to the issue requires examination of natural genetic effects on the whole of the life history, under the conditions in which it has evolved, without inducing inbreeding, and where environmental effects and parental provisioning of young do not confound genotypic effects.

We have attempted to detect additive genetic variation for net fitness by isolating the effects of a single chromosome (the third, which accounts for about 44% of the genome) in *D. melanogaster*. In this species, balancer chromosomes can be used to measure the effects of intact wild-type chromosomes. Each balancer carries inversions that suppress recombination, and a dominant marker that is also recessive lethal. In several studies, the fitness of wild-type homozygotes, relative to heterozygotes, has been measured by competing a series of wild-type chromosomes against a balancer in population cages; the equilibrium frequency reflects the relative fitnesses (Sved 1971, 1975; Wilton & Sved 1979). We extended this technique to measure heterozygous effects, relative to a standard genotype. This was done by competing each wild-type chromosome against a pair of balancers, *TM*1 and *TM*2 (Lindsley & Zimm 1992). To avoid complications from the presence of wild-type homozygotes of variable fitness, we chose wild-type chromosomes that carried a recessive lethal, so that only the three heterozygous genotypes could segregate amongst adults. Since individual recessive lethals have only slight heterozygous effects on fitness (about 2% (Simmons & Crow 1977), choosing chromosomes that happen to carry recessive lethals is unlikely to have any significant effect on the mean or variance of our fitness measure).

If the three heterozygous genotypes settled to an equilibrium, maintained by heterozygote advantage, then their relative frequencies would reflect their

* Author for correspondence. [1] Present addresses: Galton Laboratory, Department of Biology, University College London, Wolfson House, 4 Stephenson Way, London, NW1 2HE, UK. ‡ Department of Biology, Natural Science Building, University of Michigan, Ann Arbor, MI 48109, USA.

© 1997 The Royal Society

relative fitnesses. However, a pilot study had shown that the fitnesses of the three viable genotypes ranked as $TM2/+ > TM1/+ > TM1/TM2$, and that these fitnesses differed enough that cages tended to lose the $TM1$ chromosome, rather than settling to a balanced polymorphism. We therefore measured the rate at which the $TM2$ chromosome increased after a small sample of $TM2/+$ flies were introduced into cages containing only $TM1/+$, at their carrying capacity. The standard reference $TM1/TM2$ genotype arose in every experimental population cage; it should become established more rapidly and remain present for longer if the wild-type chromosome in that cage is of low fitness.

We examined the effects on net fitness of a sample of 12 third chromosomes extracted from the Dahomey base stock; each chromosome was measured in two replicate cages. The Dahomey stock has been held in population cages since 1970, some 480 fly generations, with overlapping generations and at its carrying capacity. Fitness measurements were made in cages maintained under the same regime, which can be regarded as natural conditions for this population: adaptation to laboratory culture is rapid (Briscoe *et al.* 1992; Frankham & Loebel 1992; Latter & Mulley 1995). Our results revealed strong and replicable effects of each wild-type chromosome, and indicated significant heritable variance in fitness.

2. MATERIALS AND METHODS

Stocks and cages were maintained at 25 °C on a 12 h:12 h light:dark cycle. Cages held 12 pots of sugar-yeast food medium replaced on a 4 week cycle. The Dahomey stock is of M cytotype. Third chromosomes were extracted from it using an M balancer stock $TM1/TM2$. $TM1$ and $TM2$ carry multiple inversions which suppress recombination; $TM1$ is marked with *Moiré eye* (*Me*) and $TM2$ with *Ultrabithorax* (*Ubx*) (Lindsley & Zimm 1992). The X and second chromosomes in the balancer stock were derived from Dahomey base stock. The balancer stock was regularly back-crossed to the Dahomey base stock, and was maintained as a very large (several thousand) population to ensure a normally diverse genetic background. Of the 150 third chromosomes we extracted, 30 were carrying at least one recessive lethal allele (similar to the frequency in nature (Simmons & Crow 1977)). We measured the effects on net fitness of a random sample of 12 of these. For each replicate experimental chromosome, we set up two population cages of pure $TM1/+$ genotype and two of $TM2/+$ genotype. This replication controls for accumulation of new mutations on the + chromosome, for the effects of any recombination with the balancers and for genetic drift. Numbers in the cages reached their carrying capacity after around 8–9 weeks.

We took egg samples from each of the $TM2/+$ cages by allowing adults to lay in eight vials, which were then cultured at low larval density. Adults were collected as virgins from these cultures and a group of 20 males and 20 females was added to each of the corresponding $TM1/+$ cages. The $TM2/+$ flies were introduced simultaneously into all 24 experimental cages. We monitored the subsequent change of genotypic frequencies. On three days of each week after the day of invasion, we obtained a sample of eggs from each cage by placing two pots of food medium in the cage for 1–2 h, and reared the egg samples at low larval density to minimize larval mortality. We ran the experiment for 43 weeks after the day of invasion. A total of more than 1.2×10^6 flies were counted and classified by genotype during the course of the experiment.

To estimate genotypic fitnesses from these data, we fitted a model by maximum likelihood, assuming random mating, discrete generations and constant fitnesses. Selection acted on viability differences among the three heterozygous genotypes, and was assumed to be the same for males and females. Fertility selection would be equivalent, provided there were no non-multiplicative interactions between the values for males and females (i.e. provided the relative fertility of males did not depend upon the genotype of the female and vice versa). The recursion for the frequency of the + chromosome is:

$$p_+^* = p_+ \frac{(p_{TM1} W_{TM1/+} + p_{TM2} W_{TM2/+})}{\bar{W}}, \tag{1a}$$

where

$$\bar{W} = 2(p_{TM1} p_{TM2} W_{TM1/TM2} + p_{TM1} p_+ W_{TM1/+} + p_{TM2} p_+ W_{TM2/+}). \tag{1b}$$

The equations for the frequencies of the balancer chromosomes are similar.

The parameters fitted were the initial frequency of $TM2$ (p_0), the relative fitnesses ($W_{TM1/+}/W_{TM1/TM2}$ and $W_{TM2/+}/W_{TM1/TM2}$), and the relative viabilities in the sample vials ($V_{TM1/+}/V_{TM1/TM2}$ and $V_{TM2/+}/V_{TM1/TM2}$). (These viabilities can be estimated from the proportions emerging from the vials, since at the initial and final stages of the invasion, the two genotypes carrying the rarest chromosome are at equal proportions in zygotes.)

The likelihood of each parameter combination was calculated by comparing the observed numbers of the three genotypes with the theoretical prediction, interpolated from the discrete model; binomial sampling error was assumed. Twice the difference in log likelihood between this model and a perfect fit, $(2 \log_e(L))$ has approximately a χ^2 distribution. Eighty-nine samples were taken through the experiment, each yielding two degrees of freedom; we fitted five parameters, and so the residual χ^2 was associated with $((2 \times 89) - 5) = 173$ d.f.

Residual deviations from the model were analysed by taking the difference in arcsin transformed frequency ($z = 4 \arcsin(\sqrt{p})$) between observed and predicted genotype frequencies. The variance of z due to sampling is $4/N$, where N is the number of individuals sampled. These could be compared across cages either by comparing samples at the same time, or at times when the frequency of the relevant genotype was the same. For each method, the overall mean residual deviation (which reflects deviations from the model common to all cages) was subtracted. The correlation between replicate pairs was then calculated, to reveal any residual deviations associated with the 12 chromosome lines. Comparing residuals by the second method is not straightforward: though all samples were taken on the same day, this does not correspond to the same predicted frequency. Comparison was made by taking the average of samples within ± 0.1 units of $\log_e(p/q)$, and then constructing an interpolation function using a cubic spline on this smoothed data set. A logit (i.e. $\log_e(p/q)$) scale was used to ensure approximately even spacing of sample points. Line 12 was excluded from all alignments by predicted frequency, because it did not span the same frequency range as the other lines.

In all the statistical analyses, we work with $\log_e$(fitness); correspondingly, we give arithmetic means of $\log_e$(fitness), or geometric means of fitness. We make this choice primarily

because we believe that effects are more likely to be multiplicative than additive. A logarithmic scale is also more suited to values which are necessarily positive, and which span a wide range: in principle, the fitness of $TM1/+$ and $TM2/+$ might be much smaller, or much larger, than the fitness of the standard genotype $TM1/TM2$.

3. RESULTS

The rates of change in genotypic frequencies were very different among cages containing different + chromosomes, and varied consistently across replicates. The evolution of the three most extreme lines (5, 10, 12) is shown in figure 1. In line 5, $TM2$ invaded very quickly, and became more frequent than $TM1$ after *ca.* 60 days (figure 1*a*). In line 10, $TM2$ invaded much more slowly, and only surpassed $TM1$ after 160 days (figure 1*b*). In both these extreme cases, the two replicates matched closely. In both replicate cages of line 12, $TM2$ invaded slowly; in replicate A, it was still increasing exponentially at the end of the experiment, whilst in replicate B, it appeared to level out at a low frequency of *ca.* 10% (figure 1*c*). Line 12 thus behaved quite differently from all the other lines, in which $TM1$ had declined to low frequency by the end of the experiment.

Table 1 summarizes maximum likelihood estimates of the fitnesses in the cages ($W_{TM1/+}$, $W_{TM2/+}$), the viabilities in the sample vials ($V_{TM1/+}$, $V_{TM2/+}$), and the initial frequency of $TM2$. All fitnesses and viabilities are measured relative to the standard genotype $TM1/TM2$. It is important to realize that the fitness estimates depend on the assumed generation time. Roughly speaking, the observed rates of increase per day can be fitted by assuming small fitness ratios per generation, and a short generation time, or large fitness ratios and a long generation time. The generation time is likely to lie between 15 and 25 days; we estimated fitnesses for both extremes, and found that in every case the fit was better for a shorter generation time. All our estimates therefore assume a 15 day generation time. This implies that adults reproduce mainly in the first few days of their lives, and that it is mainly the eggs which are first laid into fresh food pots that survive.

Because cage 12 A had not moved beyond the initial phase of exponential increase by the end of the experiment, it was impossible to estimate all five parameters separately for it. We could estimate only the initial rate of increase of $TM2$ from low frequency, which depends on the ratio $(W_{TM2/+}+W_{TM1/TM2})/W_{TM1/+}$, and the initial frequency of $TM2$: these estimates were 1.22 and 0.000068, respectively. Cage 12 A is excluded from the statistical analysis of fitness variation below. The data for cage 12 B could be fitted by the discrete generation model only by assuming that $TM1/+$ had high viability relative to $TM2/+$ and $TM1/TM2$ (table 1). (This is because heterozygote advantage cannot maintain a polymorphism at extreme frequency amongst zygotes; the low observed frequency of $TM2$ was therefore accounted for by low viability in the sample vials.)

Though the model fitted reasonably well, the fluctuations around the fitted curve were in all cases greater than expected from sampling error (see $\log_e(L)$ values in table 1, which are associated with 173 d.f.). This model should be seen as an approximation to the actual situation, which involves overlapping generations and age structure, and which would be described by many more parameters. However, we have simulated a model with overlapping generations,

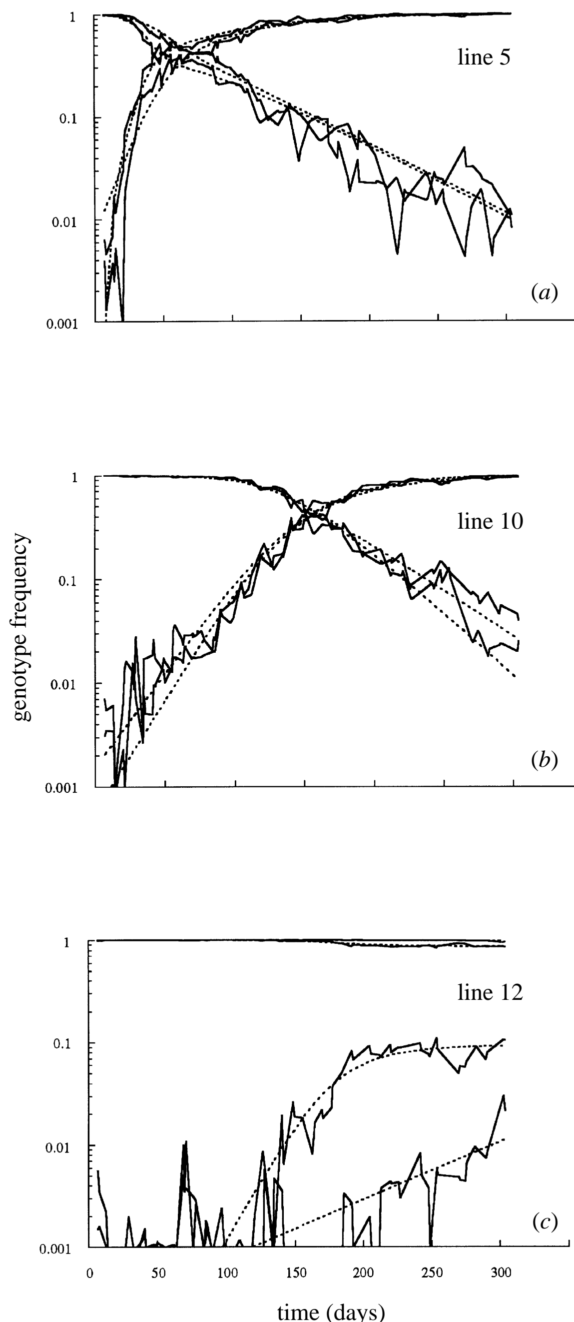

Figure 1. Changes in genotype frequencies in replicated cages plotted against time (days) since invasion. Each graph shows the observed increase in frequency of $TM2/+$, and decrease of $TM1/+$ (both replicates of both genotypes plotted as solid lines), for the two replicates of each + chromosome. (*a*) Chromosome 5, (*b*) chromosome 10, (*c*) chromosome 12. The frequency change of the $TM1/TM2$ genotype has been omitted for clarity. Chromosomes 5 and 10 showed the most extreme fitness estimates, while for chromosome 12, $TM2$ invaded very late in both replicates. The data are compared with theoretical predictions (dotted lines; parameter estimates from table 1).

Table 1. *Maximum likelihood estimates of fitness and viability for each cage* (All means are geometric means.)

chromosome	replicate	fitness		viability		initial frequency of *TM2*	$\log_e(L)$
		TM1/+	*TM2/+*	*TM1/+*	*TM2/+*		
1	A	2.958	4.740	1.273	1.192	0.000586	−462.71
	B	1.684	3.062	1.401	1.873	0.000177	−477.58
	mean	2.232	3.810	1.335	1.494	0.000322	
2	A	2.759	3.857	0.927	1.883	0.001240	−498.62
	B	2.201	3.374	0.979	2.953	0.000516	−707.69
	mean	2.464	3.607	0.953	2.358	0.000800	
3	A	2.003	4.228	1.405	1.356	0.000529	−554.05
	B	1.827	3.007	1.238	1.291	0.000608	−359.84
	mean	1.913	3.566	1.319	1.323	0.000567	
4	A	2.630	4.205	0.829	2.725	0.000441	−522.20
	B	2.580	4.093	1.061	1.790	0.000827	−361.30
	mean	2.605	4.149	0.938	2.209	0.000604	
5	A	1.107	2.667	1.684	1.178	0.005040	−962.97
	B	0.281	1.638	2.115	1.249	0.001650	−611.70
	mean	0.558	2.090	1.887	1.213	0.002884	
6	A	3.185	4.114	0.956	2.093	0.001870	−575.10
	B	2.868	4.116	1.051	1.831	0.001390	−367.00
	mean	3.022	4.115	1.002	1.958	0.001612	
7	A	1.788	2.803	1.305	1.807	0.001900	−1115.77
	B	3.245	4.789	0.921	1.626	0.001390	−366.64
	mean	2.409	3.664	1.096	1.714	0.001625	
8	A	2.438	4.247	1.224	1.741	0.000212	−365.06
	B	2.700	4.150	0.974	2.192	0.000608	−564.89
	mean	2.566	4.198	1.092	1.954	0.000359	
9	A	2.601	4.207	0.968	1.619	0.000143	−507.87
	B	1.876	3.101	1.106	1.732	0.000094	−323.49
	mean	2.209	3.612	1.035	1.675	0.000116	
10	A	7.113	12.356	1.425	2.265	0.000480	−467.68
	B	3.219	5.783	1.930	1.598	0.000343	−448.26
	mean	4.785	8.453	1.658	1.902	0.000406	
11	A	2.149	2.964	0.801	2.469	0.000310	−408.64
	B	3.103	4.537	1.018	1.574	0.000507	−255.98
	mean	2.582	3.667	0.903	1.971	0.000396	
12	A	—	—	—	—	—	—
	B	0.984	0.934	17.087	1.771	0.000120	−157.41
	mean	0.984	0.934	17.087	1.771	0.000120	
overall	mean	2.107	3.447	1.461	1.764	0.000534	

in which selection acted on either larval viability, adult longevity or adult fertility. This gave similar trajectories to the discrete model. We consider the residual deviations from the model below.

Figure 2*a* shows the estimated fitness of *TM1/+* and *TM2/+* for each cage. The two cages with highest wild-type chromosome fitness against both balancers both carried chromosome 10, whilst the two with the lowest fitness against both balancers both carried chromosome 5. The chance that the two highest, and the two lowest, pairs are replicates is $1/(21 \times 19) = 0.25\%$. The significance of the variation in fitness effects of the wild-type chromosomes is further supported by the consistency between cages 12A, 12B (figure 1*c*), and by the correlations between fluctuations around the fitted model (see below).

Figure 2*b* shows the relation between fitness and viability across cages. Fitnesses are those in the cages, whilst viabilities are those in the sample vials; both are averages across the two balancers (*TM1/+* and *TM2/+*). The viabilities vary much less than the fitnesses. Excluding line 12, the fitnesses range from 0.68 to 9.38, whereas the viabilities only range from 1.22 to 1.80. (These are averages across *TM1/+* and *TM2/+*, relative to *TM1/TM2*.) There is no correlation between average fitness and average viability ($r = 0.001$ across 22 cages, excluding line 12), and no significant variation in viability across lines ($F_{10,22} = 0.85$, $P = 0.59$). The viabilities of the *TM1/+* and *TM2/+* genotypes differed in the same direction as did their net fitnesses, but the difference was slightly smaller: the ratio of geometric mean viabilities was $V_{TM2/+}/V_{TM1/+} = 1.76/1.17 = 1.51$, compared with a ratio of average fitnesses of $3.88/2.26 = 1.72$ (excluding line 12). The narrower range of the viabilities may reflect the uncrowded conditions in the sampling vials (Kondrashov & Houle 1994); indeed, these conditions were chosen so as to minimize differential mortality.

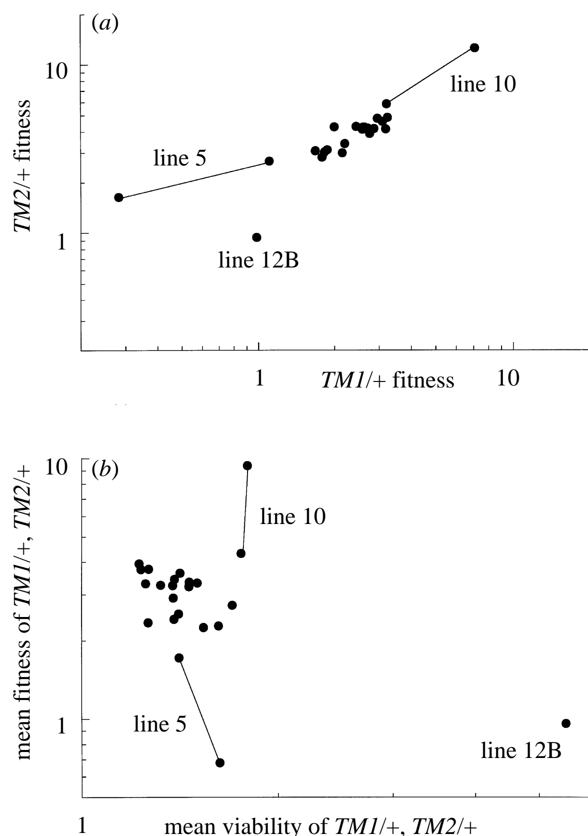

Figure 2. (*a*) Comparison between the fitness of each + chromosome in combination with $TM1$ and $TM2$ (i.e. $W_{TM1/+}/W_{TM1/TM2}$ *vs* $W_{TM2/+}/W_{TM1/TM2}$). The extreme lines (5, 10, 12) are indicated; for the first two of these, the replicate pairs are joined by lines. (*b*) Comparison between fitnesses and viabilities, for each replicate. Values are geometric means for $TM1/+$ and $TM2/+$ (i.e. $\sqrt{W_{TM1/+}W_{TM2/+}}/W_{TM1/TM2}$ vs. $\sqrt{V_{TM1/+}V_{TM2/+}}/V_{TM1/TM2}$). Chromosome 12A is excluded (see text).

The initial frequency of the $TM2$ chromosome, p_0, varied from 0.000086 to 0.0028 (geometric mean 0.00076). The correlation of $\log(p_0)$ between replicates was 0.70 ($P = 0.01$), implying that p_0 varied significantly between lines, as a consequence of the wild-type chromosome present (see, for example, line 12, figure 1*c*). This variation may reflect differences in initial population size between the lines: lines tended to have an initial frequency inversely proportional to the estimated fitness of $TM1/+$ (best fit $p_0 = 0.0021(W_{TM1/+}/W_{TM1/TM2})^{-0.94}$, $r = 0.40$; $P = 0.11$). This explanation demands large population numbers: 40 $TM2/+$ flies were put in, implying $(40/(2 \times 0.00076)) \approx 26000$ individuals in each cage. This estimate is reasonable in view of counts of around 4000 adults in a cage (Latter & Mulley 1995), since the larval and pupal populations, from which the introduced $TM2$ chromosome is initially absent, contribute to population size.

There was a strong correlation between the effects of each + chromosome in the $TM1/+$ and in the $TM2/+$ genotype on fitness (figure 2*a*). Full analysis is complicated by the fact that sampling errors in estimates of the fitnesses of $TM1/+$ and $TM2/+$ within each replicate are correlated with each other, because both are measured relative to the frequency of $TM1/TM2$ in that replicate. Since $TM1/TM2$ is relatively unfit, both the initial increase of $TM2$, and the final decline of $TM1$, depend primarily on the ratio $W_{TM2/+}/W_{TM1/+}$ which can therefore be more accurately estimated than can the fitnesses of either of these two relative to the standard genotype $TM1/TM2$. This entanglement of the fitness estimates for $TM1/+$ and $TM2/+$ makes it hard to find the extent to which the effects of the + chromosome are consistent across the two genotypes, and hence to estimate the underlying additive variance of fitness, after subtraction of measurement error.

We deal with this statistical difficulty by treating the pair of estimates of $\log_e$(fitness) as a vector, and assuming that this vector is the sum of two normally-distributed components: a measurement error with variance-covariance matrix v, and the underlying effect of the + chromosomes, with variance-covariance matrix V. Denote the two genotypes $TM1/+$, $TM2/+$ by the indices 1, 2 respectively. Then, V_{11} is the true variance between lines of $\log_e(W_{TM1/+})$, V_{12} is the true covariance between $\log_e(W_{TM1/+})$ and $\log_e(W_{TM2/+})$, and V_{22} is the true variance of $\log_e(W_{TM2/+})$. If all variation in $\log_e$(fitness) were additive, the effects of the + chromosome would be the same against each balancer, and so $V_{11} = V_{12} = V_{22}$. On the other hand, if fitness effects varied independently across balancer genotypes, then $V_{12} = 0$. Thus, hypotheses about the degree of additivity of $\log_e$(fitness) reduce to hypotheses about the structure of the covariance matrix V. Throughout, however, we make no assumptions about the structure of the error covariance matrix, v.

Because we have excluded cage 12A, this is an unbalanced data set; estimates of V and v therefore have to be found numerically by maximum likelihood. The $\log_e$(likelihoods) associated with the various hypotheses are given in table 2. The gain in log likelihood by assuming some additive variance is 1.73 with 1 d.f.; allowing a further interaction variance gives a gain of only 0.75 with 2 d.f. If one assumes the presence of interaction variance, then the gain in $\log_e$(likelihood) by allowing additive variance is 1.75 with 1 d.f. If one uses the result that $\log_e$(likelihoods) are asymptotically distributed as $(1/2)\chi^2$, then there is no evidence of interaction variance, and marginal evidence of additive variance. However, since this result is only accurate for large samples, we have used a randomization test to see whether the additive and non-additive components of variation in $\log_e$ (fitness) are significant.

The maximum likelihood estimate of the additive variance in $\log_e$(fitness) is $V_a = 0.175$, regardless of whether or not non-additive variation is included (table 2). We have generated the distribution of the estimator V_a using simulated data. This was generated for the 11 pairs of replicates, plus one unreplicated line, in two ways: first, by generating simulated data assuming no underlying variance at all ($V = 0$; hypothesis 1), and then estimating V_a on the assumption that there is no interaction variance (hypothesis 2); second, by generating data with only interaction variance ($V_{12} = 0$; hypothesis 3), and then

Table 2. *Comparison of hypotheses on the consistency of + effects against the two balancer chromosomes*

	hypothesis		d.f.	$\log_e(L)$	V_a
1	no variation between lines	$V = 0$	3	24.85	—
2	purely additive variance	$V_{11} = V_{12} = V_{22}$	4	26.58	0.1756
3	purely interaction variance	$V_{12} = 0$	5	25.57	—
4	additive and interaction variance	V arbitrary	6	27.32	0.1755

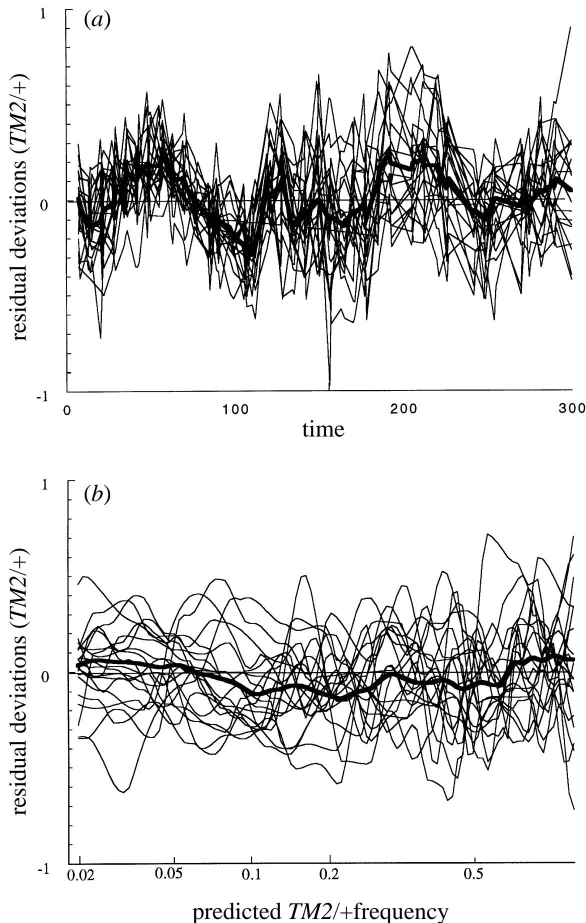

Figure 3. (a) Deviations of $TM2/+$ frequency from the fitted model, plotted against the time (in days) for each of 24 cages. (b) Deviations of $TM2/+$frequency, plotted against the predicted frequency of $TM2/+$ for each of 22 cages. $TM2/+$frequency is plotted on a logit scale. (Line 12 is excluded from the plot because it does not span the same range of frequencies as the other 11 lines.) Deviations are calculated as the difference in arcsin transformed frequency. In both graphs, the overall mean is shown as a heavy line.

estimating V_a on the assumption that there is indeed interaction variance (hypothesis 4). In each case, the data were generated using the appropriate m.l.e. for the error variance, v. Using the first method (i.e. 1 versus 2) gave $22/1000 = 2.2\%$ with V_a greater than the observed value of 0.1756. The second method (3 versus 4), allowing for an interaction effect, gave $50/1000 = 5\%$ significance. The latter test is necessarily less powerful because the presence of interaction variance across lines can generate spurious estimates of additive variance.

As a test of whether there is any significant variation between lines, we followed the geometric mean fitness, averaged across the two balancer genotypes. None of 2000 replicates gave variance estimates of V greater than the observed 0.1751. Therefore, $P < 0.15\%$. By this test, there is highly significant variation between lines. The randomization tests described above indicate a marginally significant additive component to this variance.

In every cage, deviations from the fitted model were significantly greater than expected from sampling error; the $\log_e$ (likelihood) averaged -497.50, which is 5.8 times greater than the value of 173/2 expected with 173 d.f. Figure 3a shows the residual deviations in the arcsin transformed frequency of $TM2/+$ superimposed for all 24 cages, and plotted against time ($TM1/+$ behaved in essentially the same way). There is clearly a deviation which occurs simultaneously in all cages. This could be due to common environmental perturbations which impose differential selection on the three genotypes. Alternatively, it might be due to a systematic deviation from the model. In order to distinguish these possibilities, we aligned the residual deviations by the predicted genotype frequency instead of by time. No systematic deviation is then apparent (figure 3b), indicating that deviations are caused by factors which act at the same time, rather than at the same genotype frequency. One possibility is that the fluctuations reflect differences in genotype frequency between age classes. These would be expected to fluctuate over a timescale of a few generations, which is consistent with the pattern of figure 3a. Such disequilibrium across ages would settle down from the time when the $TM2$ chromosome was introduced, rather than from the time when the invading chromosome reached appreciable frequency, and so would be consistent with the lack of overall deviation seen in figure 3b. The alternatives could be tested decisively by comparing replicate invasions of the same chromosome run simultaneously and sequentially.

Residuals peculiar to each cage can be examined by subtracting the overall mean residual. Figure 4 gives examples of these corrected residuals, for lines 5, 10 and 12, aligned by time (figures 4a–c respectively). Replicates appear to fluctuate together, indicating effects associated with different wild-type chromosomes. The similarity between replicates is striking, but is hard to test statistically because successive samples are not independent. Figure 5 shows how the correlation between replicates, across the 12 lines, changes through time. The correlation is generally positive, particularly in the middle of the experiment, when chromosome frequencies are intermediate, and sampling error should be lowest. Table 3 shows the mean correlation

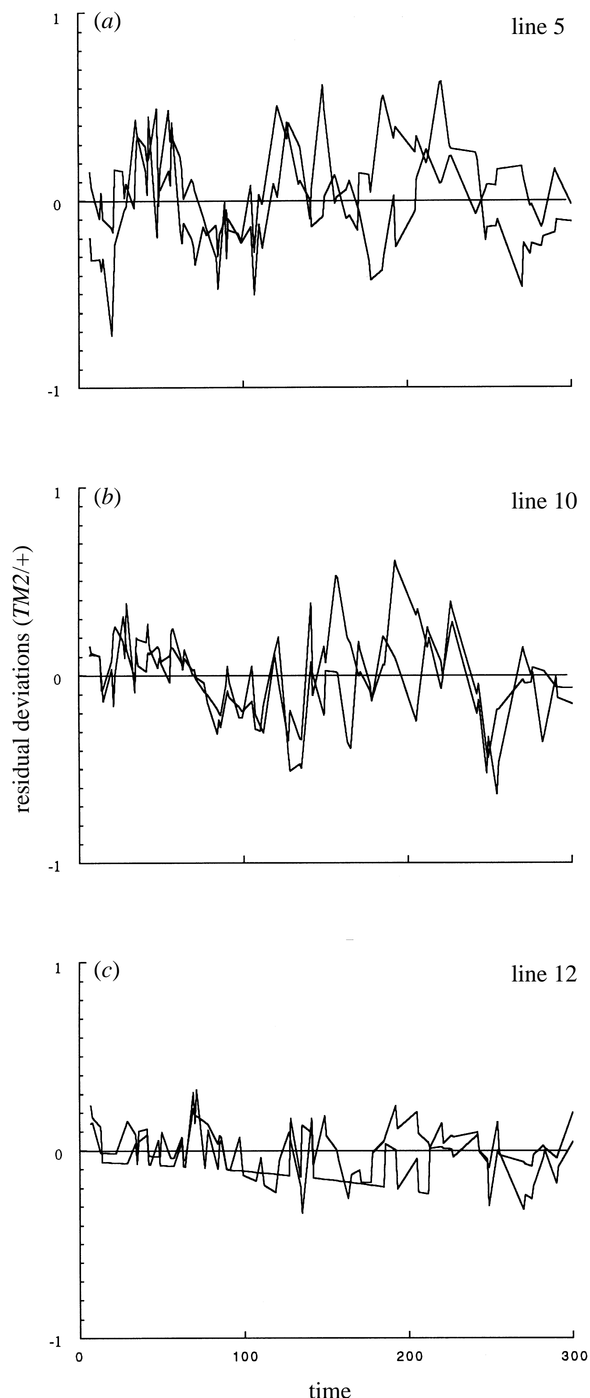

Figure 4. The residual deviation in arc-sin transformed frequency of $TM2/+$, corrected by subtraction of the overall mean (see figure 3), plotted against time (in days). (*a*) Cages 5A and 5B; (*b*) cages 10A and 10B; (*c*) cages 12A and 12B.

over the whole 300 days; the average over the two genotypes is $r = 0.167$ ($t_{11} = 2.17$, $P = 2.7\%$), suggesting a similarity between fluctuations in different cages carrying the same + chromosome. Aligning by genotype frequency gives a similar, though slightly weaker, pattern (table 3). The most plausible (indeed, the only) explanation that comes to mind is a genotype × environment interaction: that is, environmental fluctuations (for example, due to changes in food, temperature, or population density) impose different selection on different + chromosomes.

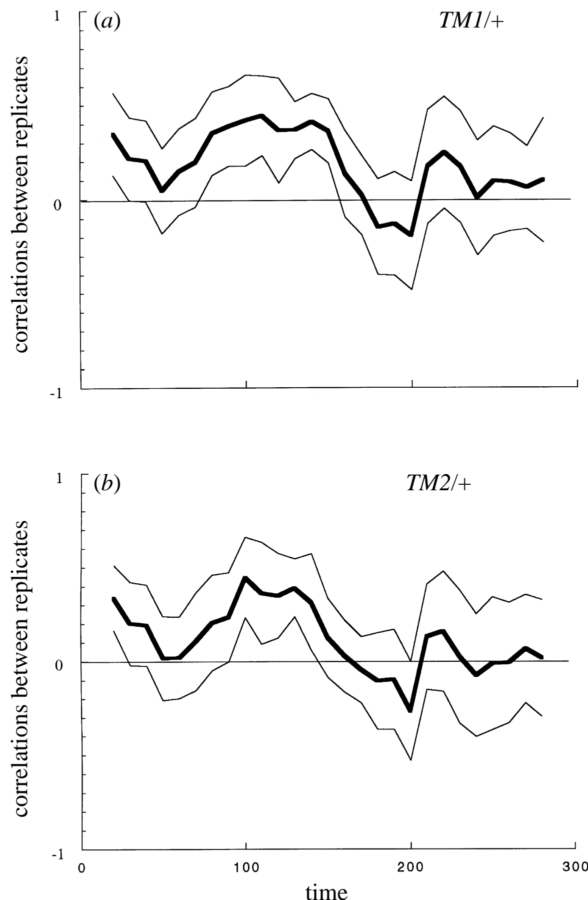

Figure 5. The correlation between replicates across the lines in corrected residuals, plotted against time (in days) for the two genotypes $TM1/+$ and $TM2/+$, respectively. The correlation is calculated using all samples within a window ± 20 days from each timepoint. The thin lines show + 2 s.e.

Table 3. *Correlations between replicates of arcsin transformed residual deviations from the fitted model*

(The data are aligned either by time, or by the predicted frequency of the relevant chromosome; in both cases, the residuals have been corrected by subtraction of the overall mean. For the data aligned by time, the correlation averages over the two chromosomes and over all lines is 0.167; $t_{11} = 2.17$, $P = 2.65\%$. For alignment by frequency, the mean correlation is 0.139; $t_{10} = 1.59$, $P = 7.11\%$.)

	aligned by time		aligned by frequency	
chromosome	*TM*1	*TM*2	*TM*1	*TM*2
1	0.191	0.065	0.111	−0.006
2	0.367	0.453	0.306	0.172
3	−0.056	−0.098	0.057	0.039
4	−0.030	−0.205	−0.052	−0.141
5	0.124	−0.077	0.602	0.652
6	0.212	0.149	−0.218	−0.237
7	−0.157	−0.320	−0.082	−0.087
8	0.046	−0.142	0.110	0.235
9	0.272	0.241	0.441	0.563
10	0.563	0.416	0.563	0.218
11	0.406	0.398	−0.070	−0.127
12	0.609	0.587	—	—
mean	0.213	0.123	0.161	0.116

4. DISCUSSION

We have demonstrated significant variation in heterozygous effects of different wild-type chromosomes on net fitness. The consistency of effects of different wild-type chromosomes in combination with the two balancers suggests that the effects we have measured are additive, and do not rely on interaction effects with the balancers. Because we have examined the effects of whole third chromosomes, it is possible that epistatic interactions between the loci on each chromosome contribute to their effects on net fitness. This possibility should be tested, and this system of measurement would allow such a test. A previous study (Charlesworth & Charlesworth 1975) has detected significant costs to female fecundity of recombination within a chromosome in this species, suggesting that such intra-chromosomal epistatic fitness effects can occur. Although they may not contribute to long-term evolutionary change, they could be important for mate choice, since any genetic effect on fitness that is not completely destroyed between one generation and the next can contribute to the success of offspring of a mating.

We measured the relative fitness effect of each wild-type chromosome by the geometric mean of $TM1/+$ and $TM2/+$ fitnesses for that line, divided by the overall geometric mean. The values span a range from 0.40 to 1.70 with a between-line variance of 0.10. This value is corrected for error variance estimated from the difference between replicates. Using an additive scale makes little difference since for these data the variance in fitness, standardized relative to the arithmetic mean, is 0.06. These figures should be regarded as provisional, given the small sample size of 12 chromosomes, and the observed distribution of fitness effects, with three chromosomes of extreme effect contributing most of the variation in fitness. A larger sample of chromosomes should be examined to obtain a more reliable distribution of effects. In addition, the analysis could be extended to chromosomes that do not happen to be carrying a recessive lethal.

Since we estimate the effect on fitness of one chromosome out of two, the above values correspond to half the genotypic variance in fitness associated with the third chromosome. The third chromosome contains 44% of the total number of polytene bands in *Drosophila* (Charlesworth *et al*. 1992), and so extrapolating to the whole diploid genome gives an estimated genotypic variance in net fitness of 0.45. Our estimate may reflect the value in nature if mutation–selection balance is the main cause of standing genetic variance for net fitness, or may be lower than the value in nature if the intensity or variability of selection is reduced under laboratory conditions (Felsenstein 1976; Charlesworth 1987; Kondrashov & Houle 1994). It is not clear if our estimate could be accounted for by mutation–selection balance alone, because previous estimates of mutational variance for traits such as morphology (e.g. Keightley *et al*. 1993) and viability (e.g. Mukai 1964; Mukai *et al*. 1972) may be much lower than the real values for net fitness, which have not been directly measured. Recent estimates suggest that mutational variance may be particularly high for fitness-related traits such as life history (Houle *et al*. 1996) and male mating success (Rowe & Houle 1996), perhaps because so many genes contribute to them, and there is a possibility that much of the high standing genetic variance for these traits can be accounted for by mutation–selection balance (Houle *et al*. 1996).

It will also be important to discover which fitness components contribute to additive genetic variance in net fitness, and the genetic correlations between different fitness components, including between the sexes. Our system can be used to examine the contributions of different fitness components, including male mating success, to net fitness, and hence to test the plausibility of good genes models of sexual selection and of theories of life history evolution.

We thank John Sved for helpful discussions in the planning stages of the project, Brian Charlesworth, Alexei Kondrashov, Trudy Mackay and Steve Stearns for comments on the manuscript, SERC, BBSRC, the Darwin Trust and the Royal Society for financial support, and Ms N. Goorney for technical assistance.

REFERENCES

Briscoe, D. A., Malpica, J. M., Robertson, A., Smith, G. J. & Frankham, R. 1992 Rapid loss of genetic variation in large captive populations of *Drosophila* flies—implications for the genetic management of captive populations *Cons. Biol*. **6**, 416–425.

Brittnacher, J. G. 1981 Genetic variation and genetic load due to the male reproductive component of fitness in *Drosophila*. *Genetics* **97**, 719–730.

Bundgaard, J. & Christiansen, F. B. 1972 Dynamics of polymorphisms. I. Selection components in an experimental population of *Drosophila melanogaster*. *Genetics* **71**, 439–460.

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

Charlesworth, B. 1987 The heritability of fitness. In *Sexual selection: testing the alternatives*, (ed. J. W. Bradbury & M. B. Anderson), pp. 21–40. Wiley.

Charlesworth, B. & Charlesworth, D. 1975 . An experiment on recombination load in *Drosophila melanogaster*. *Genet. Res. Camb*. **25**, 267–274.

Charlesworth, B., Lapid, A. & Canada, D. 1992 The distribution of transposable elements within and between chromosomes in a population of *Drosophila melanogaster*. II. Inferences on the nature of selection against elements. *Genet. Res*. **60**, 103–114.

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

Felsenstein, J. 1976 The theoretical population genetics of variable selection and migration. *A. Rev. Genet*. **10**, 253–280.

Frankham, R. & Loebel, D. A. 1992 Modelling problems in conservation genetics using captive *Drosophila* populations—rapid genetic adaptation to captivity. *Zoo Biol*. **11**, 333–342.

Gustaffson, L. 1986 Lifetime reproductive success and heritability: empirical support for Fisher's fundamental theorem. *Am. Nat*. **128**, 761–764.

Haymer, D. S. & D. L. Hartl. 1981 Using frequency distributions to detect selection: inversion polymorphism in *Drosophila pseudoobscura*. *Evolution* **35**, 1243–1246.

Haymer, R. S. & Hartl, D. L. 1982 The experimental assessment of fitness in *Drosophila*. I. Comparative measures of competitive reproductive success. *Genetics* **102**, 455–466.

Houle, D., Morikawa, B. & Lynch, M. 1996 Comparing mutational variabilities. *Genetics* **143**, 1467–1483.

Hughes, K. A. 1995 The evolutionary genetics of male life-history characters in *Drosophila melanogaster*. *Evolution* **49**, 521–537.

Keightley, P. T., Mackay, T. F. C. & Caballero, A. 1993 Accounting for bias in estimates of the rate of polygenic mutation. *Proc. R. Soc. Lond.* B **253**, 291–296.

Kirkpatrick, M. & Ryan, M. J. 1991 The evolution of mating preferences and the paradox of the lek. *Nature, Lond.* **350**, 33–38.

Kondrashov, A. S. & Houle, D. 1994 Genotype–environment interactions and the estimation of the genomic mutation rate in *Drosophila melanogaster*. *Proc. R. Soc. Lond.* B **258**, 221–227.

Latter, B. D. H. & Mulley, J. C. 1995 Genetic adaptation to captivity and inbreeding depression in small laboratory populations of *Drosophila melanogaster*. *Genetics* **139**, 255–266.

Lindsley, D. L. & Zimm, G. G. 1992 . *The genome of Drosophila melanogaster*. San Diego, CA: Academic Press.

Maynard Smith, J. 1978 *The evolution of sex*. Cambridge University Press.

Mukai, T. 1964 The genetic structure of natural populations of *Drosophila melanogaster*. I. Spontaneous mutation rate of polygenes controlling viability. *Genetics* **50**, 1–19.

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**, 335–355.

Norris, K. 1993 Heritable variation in a plumage indicator of viability in male great tits *Parus major*. *Nature, Lond.* **362**, 537–539.

Partridge, L. 1980 Mate choice increases a component of offspring fitness in fruitflies. *Nature, Lond.* **283**, 290–291.

Partridge, L. 1983 Non-random mating and offspring fitness. In *Mate choice*, (ed. P. Bateson), pp. 227–255. Cambridge University Press.

Partridge, L. & Fowler, K. 1992 Direct and correlated responses to selection on age at reproduction in *Drosophila melanogaster*. *Evolution* **46**, 76–91.

Partridge, L. & Fowler, K. 1993 Responses and correlated responses to artifical selection on thorax length in *Drosophila melanogaster*. *Evolution* **47**, 213–226.

Petrie, M. 1994 Improved growth and survival of offspring of peacocks with more elaborate tails. *Nature, Lond.* **371**, 598–599.

Rose, M. R. 1984*a* Laboratory evolution of postponed senescence in *Drosophila melanogaster*. *Evolution* **38**, 1004–1010.

Rose, M. R. 1984*b* Genetic covariation in *Drosophila* life history: untangling the data. *Am. Natur.* **123**, 565–569.

Rowe, L. & Houle, D. 1996 The lek paradox and the capture of genetic variance by condition-dependent traits. *Proc. R. Soc. Lond.* B **263**, 1415–1421.

Service, P. M. & Rose, M. R. 1985 Genetic covariation among life-history components: the effect of novel environments. *Evolution* **39**, 943–945.

Simmons, M. J. & Crow, J. F. 1977 Mutations affecting fitness in *Drosophila* populations. *A. Rev. Genet.* **11**, 49–78.

Stearns, S. C. 1992 *The evolution of life histories*. Oxford University Press.

Sved, J. A. 1971 An estimate of heterosis in *Drosophila melanogaster*. *Genet. Res.* **18**, 97–105.

Sved, J. A. 1975 Fitness of third chromosome homozygotes in *Drosophila melanogaster*. *Genet. Res.* **25**, 197–200.

Williams, G. C. 1975 *Sex and evolution*. Princeton: Princeton University Press.

Wilton, A. N. & Sved, J. A. 1979 X-chromosomal heterosis in *Drosophila melanogaster*. *Genet. Res.* **34**, 303–315.

Received 28 August 1996; accepted 9 October 1996

As this paper exceeds the maximum length normally considered for publication in *Proceedings* B, the authors have agreed to make a contribution towards production costs.