Genet. Res., Camb. (1989), **53**, *pp.* 129–140 *With 4 text-figures Printed in Great Britain*

Can chromosomal heterosis in *Drosophila* be explained by deleterious recessive genes? Negative results from a dichromosomal population test

ALAN N. WILTON*, MICHAEL G. JOSEPH AND JOHN A. SVED
School of Biological Sciences A12, University of Sydney, NSW 2006, Australia

(*Received 16 August 1988*)

Summary

High levels of chromosomal heterosis have previously been detected in *Drosophila* using the balancer chromosome equilibration (BE) technique, in which single wild-type chromosomes are introduced into population cages along with a dominant/lethal balancer chromosome. The balancer chromosome is rarely eliminated in such populations, showing that the fitness of chromosome homozygotes must be low by comparison with chromosomal heterozygotes. As with all cases of chromosomal heterosis, the underlying cause could either be deleterious recessives at various loci or generalized overdominance. The experiment of the present paper examines the first of these explanations. Population cages containing just two wild-type chromosomes (dichromosomal populations) were set up and allowed to run for many generations. Single chromosomes were then re-extracted from these populations, and their fitness measured using the BE technique. Our expectation was that the gradual elimination of recessive genes from the dichromosomal populations ought to result in an increase in the fitness of such re-extracted chromosome homozygotes. Yet in two replicated experiments we were unable to demonstrate any unequivocal increase in fitness. We have estimated the rate of increase of fitness under multiple locus dominance and partial dominance models. The principal unknown parameter in these calculations is the selection intensity per locus, s. The expected increase is approximately proportional to s, and we estimate that values of s around 1/64 should be detectable in our experiments. However linkage is expected to reduce the efficiency of the dichromosomal procedure. We show by computer simulation that this reduction is by a factor of approximately 2, thus increasing the detectable level of s to approximately 1/32. Consideration of mutation-selection balance models shows that this is a feasible selection intensity only if dominance is nearly complete. Thus we are unable to rule out the notion that the genes responsible for heterosis are maintained by a simple mutation-selection balance, but the experimental results constrain the parameters of such a model to a narrow range.

1. Introduction

(i) *Heterosis in* Drosophila

The original demonstrations of heterosis in *Drosophila* date from the work of Gowen (1952) and others who measured viability and fecundity in individuals having various levels of inbreeding. These studies were made using mating between relatives, which produces unpredictable homozygosity in any given region, in addition to variable overall levels of inbreeding. The introduction of balancer chromosomes enabled such studies to be carried out in a more controlled manner on individuals homozygous for particular chromosomes (e.g. Dobzhansky & Spassky, 1954; see Lewontin, 1974 for an overall discussion of this area). These experiments indicated that chromosomal heterozygotes survive and reproduce at a significantly higher rate than homozygotes, although the selective differentials were not large.

The balancer chromosome approach was extended by Sved & Ayala (1970), who introduced the 'balancer equilibration' (BE) technique. This gives a single overall estimate of chromosome homozygote fitness by monitoring the frequency of balancer chromosomes in population cages (see also Sperlich & Karlik, 1970). This technique, with minor modifications, has been carried out on several different species and chromo-

* Current address: School of Biological Sciences, Macquarie University, NSW, Australia.

somes (see Simmons & Crow, 1977, table 4). In all cases the results have indicated that chromosome homozygotes suffer a high depression in fitness, around 80%, compared to chromosomal heterozygotes. Thus the problem of accounting for heterosis in *Drosophila* has been greatly accentuated.

(ii) *Dominance and overdominance models*

It has been known since the work of Davenport (1908), Schull (1908) and East (1908) that heterosis can be explained either in terms of overdominance (heterozygote advantage) or of dominance (deleterious recessives) at individual loci. Much of the early discussion of these concepts took place in the context of heterosis for yield in corn. Although the present paper concerns the measurement of fitness in *Drosophila*, the same distinction between dominance and overdominance models can be made here.

There seems currently to be relatively little support for the overdominance model. This is partly based on the failure to identify clearcut examples of overdominance, as opposed to numerous known examples of deleterious recessives. While the early evidence in corn was frequently interpreted in terms of overdominance (e.g. Hull, 1952), most long-term experiments have indicated that dominance or partial dominance of linked genes is a more likely explanation (e.g. Sprague, Russell & Penny, 1956; Gardner, 1963). In other organisms, the frequency distribution of electrophoretically-detectable genes shows an elevation in the incidence of genes at low frequency rather than genes at central frequencies (Yamazaki & Maruyama, 1971). This favours the notion that the majority of such genes are held at low frequency by a mutation–section balance. The finding of high levels of variability in some haploid organisms is also evidence against overdominance, at least as an explanation for maintaining variability in these cases (Kimura, 1983, p. 271).

None of these arguments bears directly on the explanation of chromosomal heterosis in *Drosophila*. This paper reports the results of a set of experiments performed to test a prediction from the dominance hypothesis. The basis for this prediction is similar to that put forward by Schull (1911) and East & Hayes (1912), who argued that under the dominance hypothesis it should be possible to select out the deleterious genes, leading to pure lines having fitness equal to the heterozygote. As pointed out by Jones (1917) and Collins (1921), there are difficulties associated with the application of this test, resulting from either linkage or the numbers of loci involved. These points are considered in some detail in a later section in which we calculate the power of this prediction in the *Drosophila* system.

(iii) *The dichromosomal experiment*

The outline of the procedure is shown in Fig. 1. It shows that two individuals, or equivalently two single chromosome lines, are intercrossed to produce a population of heterozygous individuals. It is these individuals which found the 'dichromosomal' populations.

Figure 1 shows that there are two types of population cages. (1) Balancer equilibration cages used for measuring homozygous fitness, shown with light tops in the figure. We will refer to these for brevity as [B+] populations throughout the paper. (2) Dichromosomal cages, which are long-term cages, up to 2 years, shown in the figure with dark tops. These

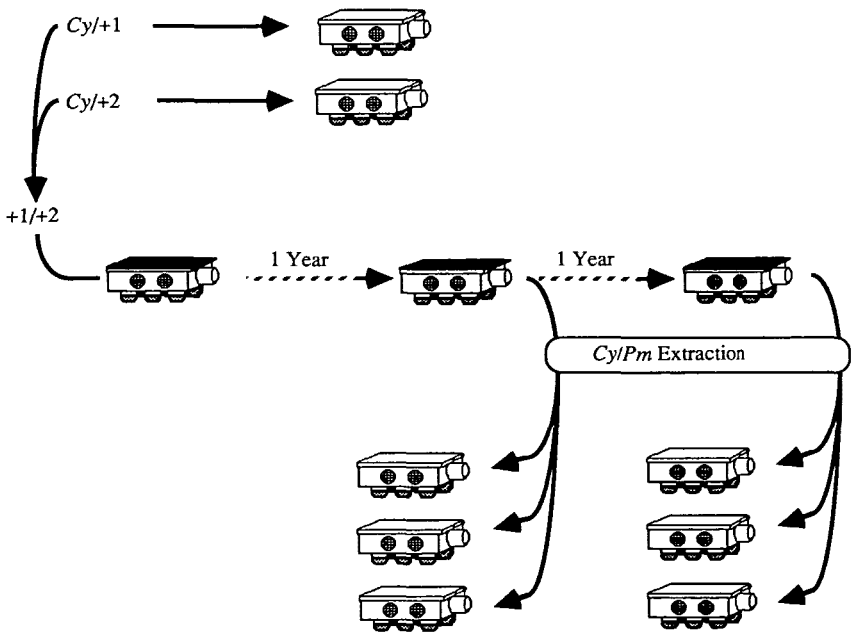

Fig. 1. Outline of the dichromosomal cage test.

https://doi.org/10.1017/S0016672300028019 Published online by Cambridge University Press

will be referred to as [+ +] populations, signifying that they contain two wild-type chromosomes and no balancer chromosome.

A third type of cage which does not appear in Fig. 1 but which was used in two different places in these experiments is the 'heterozygous' cage, in which the balancer chromosome is present along with a mixture of wild-type chromosomes.

The dichromosomal procedure has the effect of initiating a population in which all deleterious genes are present at frequency 1/2. Assuming that deleterious genes are individually rare in the wild population, each such gene should initially be present on only one of the two founder chromosomes. Selection against a deleterious gene is very inefficient when the gene is rare. Such selection is maximized at a frequency of approximately 2/3 for a recessive gene and 1/2 for a non-recessive gene. Thus the opportunity for selection should be nearly maximized in the [+ +] cages.

The results of the experiment come from the chromosome frequencies in the [B+] populations. If chromosomal heterosis is attributable to deleterious genes, then as these genes are eliminated from the [+ +] populations the fitness of homozygotes of chromosomes extracted from these populations should rise. The end-point of this process would presumably be the elimination of all deleterious genes, in which case the fitnesses of chromosome homozygotes and heterozygotes would be identical. Our aims in these experiments were therefore twofold: (1) to test for a rise in the mean fitness of chromosome homozygotes over the time course of the [+ +] populations; (2) to test for the production of 'superfit' chromosome homozygotes.

2. Methods and materials

(i) *Stocks*

Wild type. Stocks were collected from wineries in the Hunter Valley district of NSW just before use in the experiments.

SM1(CY)/Pm. This stock was re-synthesized by crossing to a mixed Hunter Valley wild-type stock, intercrossing, and then maintaining in bottles for many generations. The stock is classified as a P stock in the P–M hybrid dysgenesis system (Kidwell, Kidwell & Sved, 1977), which avoids mutational effects when the crosses are made in the chromosome extraction process.

SM1(Cy)/Pm; pol. The fourth chromosome marker spa^{pol} was added to the *SM1(Cy)/Pm* stock above. This stock is also classified as P in the P–M system.

Crosses. Single chromosomes for the experiment were extracted using the standard *Cy/Pm* method (see Sved, 1971). The crossing procedure used two extra generations of crossing *en masse* to the above *Cy/Pm* stock in order to minimize the inbreeding on chromosomes other than II. The initial cross involving wild-type males would be expected to lead to dysgenesis if a standard laboratory stock was used, but the introduction of wild-type chromosomes into the *Cy/Pm* stock should alleviate this problem.

Chromosomes for the dichromosomal cages were chosen after the results of viability tests were available. Chromosomes having lethal genes or viability less than 50% were excluded. Owing to the length of time taken to obtain overall chromosome fitness estimates, only viability estimates were available at the time the dichromosomal cages were set up. A potential problem is caused where the two founder chromosomes differ markedly in fitness, since segregation of chromosome regions can be expected to raise the average fitness in the population to that of the higher chromosome fitness. This process can occur without any overall decrease in the frequency of deleterious recessives which the dichromosomal experiment is set up to detect. To some extent this problem can be overcome by taking an early sample from the [+ +] population. Later samples can then be compared to this first sample rather than to the estimates associated with the founder chromosomes.

It should be noted that three 'sampling' processes are involved in this experiment. The first involves the sampling of a number of wild type chromosomes to initiate the [+ +] populations, using the *Cy/Pm* procedure. The second involves sampling a number of chromosomes from the [+ +] population, again using the *Cy/Pm* procedure. The third involves counting a sample of progeny from the [B+] cages to establish chromosome frequencies. Where there is some chance of confusion, particularly between the second and third processes, we will refer to the former as 'chromosome sampling', or sometimes 'chromosome extraction', and the latter as 'population sampling'.

(ii) *Experiment 1*

We report the results of two experiments in this paper, both of which use the basic design given in Fig. 1. The two experiments differ in several ways, and were carried out at different times. The timing of the first experiment is shown in tabular form (Table 1), and of the second experiment in diagrammatic form (Fig. 2). Table 1 shows that four [+ +] populations were set up, three in duplicate. Up to four times during the lifetimes of these cages, chromosomes were extracted for testing in [B+] populations. In this experiment the founding chromosome [B+] populations were kept and repeatedly sampled throughout the period in which the [+ +] populations were maintained. Detailed results for this experiment are given in Wilton (1980).

The chromosomes for the first experiment were taken from a sample of 24 single-chromosome lines extracted from Hunter Valley populations. Ten of these lines were rejected as having lethals, or low or

https://doi.org/10.1017/S0016672300028019 Published online by Cambridge University Press

Table 1. *Founder chromosomes, their BE frequencies, [+ +] populations, sampling times in weeks and number of sample cages for experiment 1*

+/+ (1)	+/+ (2)	Line	Sampling time (no. of cages)
0·409 (m.s.)	0·397	A1	11 (8), 22 (4), 44 (8), 77 (16)
		A2	12 (4), 72 (8)
0·389	0·447	B1	11 (7), 22 (6), 44 (6), 77 (18)
		B2	11 (20), 72 (5)
0·324 (f.s.)	0·495	C1	11 (6), 22 (6), 44 (8), 77 (20)
		C2	11 (6), 72 (4)
0·367	0·523	D	12 (7), 72 (10)

variable viability. The eight founding chromosomes were chosen from the remainder, and included one with a male sterile gene and one with a female sterile gene, as shown in Table 1. These were retained in the experiment to allow testing for the effect of major fitness genes.

When chromosomes were sampled in the first experiment, as well as setting up the individual [B+] sampling cages, we also set up cages which contained a mixture of the different [B+] cages. These may be described as 'heterozygous sample populations', since wild-type flies in these cages are heterozygous rather than homozygous. Our expectation was that the balancer chromosome would be eliminated from these cages, and that the reduction in frequency of deleterious genes from the [+ +] populations would lead to an increased rate of elimination of the balancer chromosome throughout the course of the experiment. Since the +/+ homozygotes in this test consist of a mixture of genotypes from two original chromosomes, the test is not as sensitive as those involving a single chromosome.

The experiment of Sperlich & Karlik (1970) should be mentioned in this context. These authors constructed cages containing one, two (dichromosomal), or more wild-type chromosomes, in addition to a balancer chromosome. The theory behind the use of the dichromosomal population is similar to that of our experiment, except that the processes of selection against deleterious genes and selection against the balancer chromosome are telescoped into a single population. The design of Fig. 1 allows the process of change of chromosome fitness to be monitored more directly over a much longer period of time.

Subsequent to setting up the [+ +] populations in the first experiment, it was discovered that all had inadvertently been set up with one of the two founder chromosomes containing an inversion in 2*L*. Since recombination was therefore expected at a low level in this chromosome arm, the experiment was essentially restricted to a test for loss of recessives on only half of the chromosome, and necessitated the running of a second experiment.

(iii) *Experiment 2*

The timing of extractions and sampling is illustrated in Fig. 2. All chromosomes containing sterile genes, as well as lethals and low viability genes, were rejected for this experiment. The other important difference

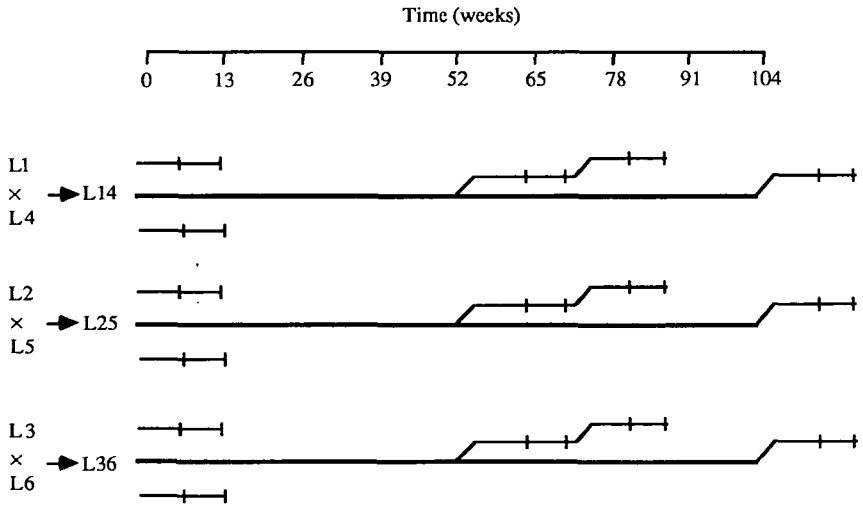

Fig. 2. Approximate time course of experiment 2. Thick horizontal lines represent the dichromosomal [+ +] populations, thin horizontal lines represent balancer equilibration [B+] populations, diagonal lines represent the chromosome sampling process, and verticle lines represent population sampling. Replication is not shown, although each [+ +] population was duplicated and each chromosome sampling was replicated many times.

https://doi.org/10.1017/S0016672300028019 Published online by Cambridge University Press

between experiments is that for the second experiment, the marker spa^{pol} on chromosome IV was introduced into the dichromosomal populations, to guard against the possibility of contamination of the cages during the long time period of the experiment.

(iv) *Retesting of chromosome fitness*

One possible outcome of the experiments is the production of 'superfit' chromosome homozygotes, in which case the balancer chromosome will be eliminated from a [B+] population. Unfortunately this result is not very different to what is expected from a population which is inadvertently initiated with a mixture of wild type chromosomes. However a simple re-test is possible which should unequivocally distinguish between these two possibilities. A series of [B+] populations can be initiated by re-sampling single chromosomes from the [B+] cage in question using the *Cy/Pm* technique. In the event that the original result is due to a superfit chromosome, *all* derivative populations should again eliminate the balancer chromosome. If the result is due to a mixture of chromosomes, derivative cages should revert to a low chromosome fitness and retain the balancer chromosome.

(v) *The use of viability tests*

Estimation of the overall chromosome fitness requires the use of chromosome frequencies at the zygote stage (see e.g. Prout, 1965). Such frequencies cannot be studied directly. However use of a controlled cross involving the same chromosomal genotypes, grown under similar conditions to the population sample, enables the zygote frequencies to be inferred.

Based on the results of viability crosses from several studies in which homozygous fitness has been estimated, there appears to be a bias associated with the procedure. The most striking manifestation of this bias comes from the negative fitness estimates obtained for some chromosomes (e.g. Sved, 1971). Such cases arise when the frequency of $+/+$ genotypes coming from a $Cy/+ \times Cy/+$ cross is higher in the frequency sampled from the population cage. Since the contribution from $+/+$ in the cage population cannot be less than zero, any such significantly negative difference must be attributable to some difference in conditions between the cage and viability sample. Because of the high variance of the fitness estimates, it has been difficult to draw conclusions about the significance of these negative estimates. This question has been investigated in more detail by Pascoe (1985). Preliminary results have indicated, surprisingly, that the manner in which the parents are raised may have an effect on the distribution of genotypes amongst their progeny. Thus the frequency of $+/+$ in $Cy/+ \times Cy/+$ crosses may be lower if these parents are cage reared than if they are reared in bottles under conditions of low crowding such as usually done in virgin-collecting procedures. A strictly valid viability estimate would thus require that the parents used in a viability cross be cage-reared. The collection of virgin females for a viability cross then poses special problems.

In view of the potential bias, the extra stochastic variation introduced, and the problems involved in carrying out the viability tests, there are clear advantages in eliminating this part of the procedure altogether, a course of action advocated by Haymer & Hartl (1982). This is particularly the case in the present experiment where we are testing for a rise in the fitness of the $+/+$ homozygote. Such a rise can be detected directly by a rise in the equilibrium frequency of the $+/+$ genotype in samples taken from the dichromosomal populations. To convert this frequency into a fitness estimate requires the viability sample. However if our primary interest is in finding whether there has been *any* significant rise in fitness, then it seems that this question is best studied without bringing in the complication of the viability test. Also if the initial chromosomes are chosen as having viability not significantly different from unity, then no large change in viability would be expected over the course of the experiment.

(vi) *Homozygous and heterozygous populations*

The overall estimate of homozygous fitness from [B+] populations is obtained relative to the $Cy/+$ genotype. In order to estimate the fitness of chromosome homozygotes relative to chromosome heterozygotes we must use the results from [B+] cages set up with mixtures of wild type chromosomes. These give a maximum estimate of fitness of about 50% for the $Cy/+$ genotype relative to $+/+$ heterozygotes (Sved, 1971). Based on this figure, the estimate of $+/+$ homozygous fitness compared to $Cy/+$ is halved to give the estimate of $+/+$ homozygote fitness compared to $+/+$ heterozygotes.

Fortunately, the complication arising from this dual estimation process does not affect the present experiment. In this case the selective differentials of $+/+$ homozygotes to $+/+$ heterozygotes apply directly in the [++] populations. The $Cy/+$ genotype is not present and is not relevant to the processes determining the rate of loss of deleterious genes.

The estimation of fitnesses is, however, made relative to the $Cy/+$ genotype. In fact this considerably increases the sensitivity of the experiment. This can be seen by considering the consequences of a 50% increase in the fitness of the $+/+$ homozygote. If this was measured directly against the $+/+$ heterozygote, we would expect the fitness to rise from 20 to 30%. However, measured against the less fit $Cy/+$ heterozygote, the fitness is expected to rise from 40 to 60%, a more substantial and easily detectable rise.

https://doi.org/10.1017/S0016672300028019 Published online by Cambridge University Press

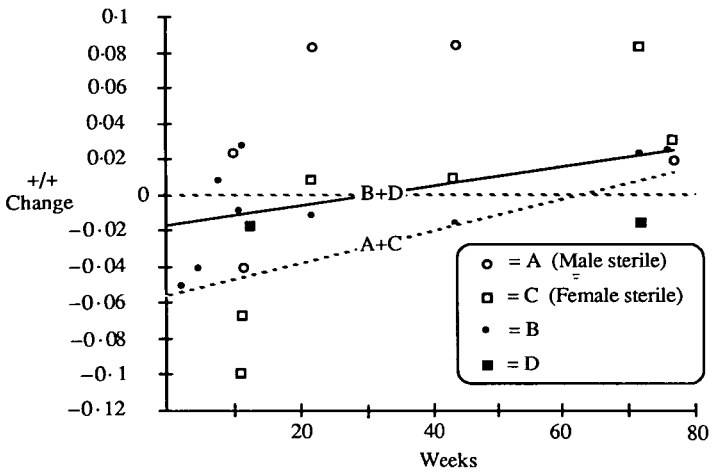

Fig. 3. Estimated change and regression lines for frequency of the +/+ genotype in chromosome samples taken at various times in experiment 1.

(vii) *Computer simulation*

Two computer programs were written to simulate the expected outcome of the dichromosomal experiment under a range of parameters of the deleterious recessive model. These programs were written respectively in Fortran for the CDC Cyber computer and in C for the Macintosh computer. Listings of the programs are available on request. The fact that similar conclusions were produced by two independently written programs gives us added confidence in the validity of the results. Both programs work in a strictly Monte-Carlo fashion by simulating a diploid population of finite size, with each gene being separately represented. In both cases we used individual bits to represent genes, thereby allowing many linked loci (60 per word on the Cyber, and 16 per short integer word on the Macintosh) to be represented in a single word. Independence of loci could also be simulated by using randomly generated masks.

3. Results

(i) *Experiment 1*

Fig. 3 shows the change in the frequency over the course of the experiment of the +/+ genotype, as determined from the [B+] sample populations. All frequencies are expressed relative to the mean frequency from the [B+] founder chromosome populations. Data points in the figure are distinguished according to whether they come from one of the two populations started with a sterility gene (open symbols) or not (closed symbols). Each data point is based on the average of all chromosomes sampled at that time (see Table 1), although omitting chromosomes found to have newly produced lethal mutations. There are clearly much wider fluctuations in the cases where a major fitness gene is involved in the [++] population. The regression of fitness on time is greater in the former case ($b_{A+C} = -0{\cdot}0081$, $t = 2{\cdot}6$, $P < 0{\cdot}05$; $b_{B+D} = -0{\cdot}0048$, $t = 2{\cdot}2$, $P < 0{\cdot}05$), as would be expected from loss of the sterility genes. However in neither case is there any large increase in the frequency of the +/+ homozygous genotype such as would be predicted under the deleterious gene hypothesis.

Heterozygous cages were set up at most samplings in this experiment to test for the rate of elimination of the balancer chromosome. Increases in +/+ fitness were found, but only for the later samples (77 weeks) of the two populations set up with sterility genes. While not as sensitive a test as that provided by the homozygous cages, nevertheless the results from these populations confirm the absence of a high rate of loss of deleterious genes from the wild-type chromosomes.

(ii) *Experiment 2*

As indicated in Fig. 2, samples in this experiment were taken after one and two years. Because of problems in interpreting the results of the one-year sample, we have omitted this from the analysis. The sample initially showed highly significant increases in the frequency of the +/+ genotype for 10 out of 24 sampled chromosomes. Viability tests were also made for all chromosomes, and 4 out of the 10 chromosomes with ostensibly high fitness were found to have reductions in viability of more than 50%, clearly contradicting the fitness estimates. Furthermore we re-tested the fitnesses by re-sampling from the cages as indicated in the Methods and Materials section. In almost all cases the chromosomes failed to reproduce the results from the earlier cages. In view of all the inconsistencies it seems most appropriate to ignore the results from this sample and to concentrate on the results from the later sample, in which no contradictory results were obtained.

The anomolous results from the first isolation led us to adopt a more concentrated sampling in the second

https://doi.org/10.1017/S0016672300028019 Published online by Cambridge University Press

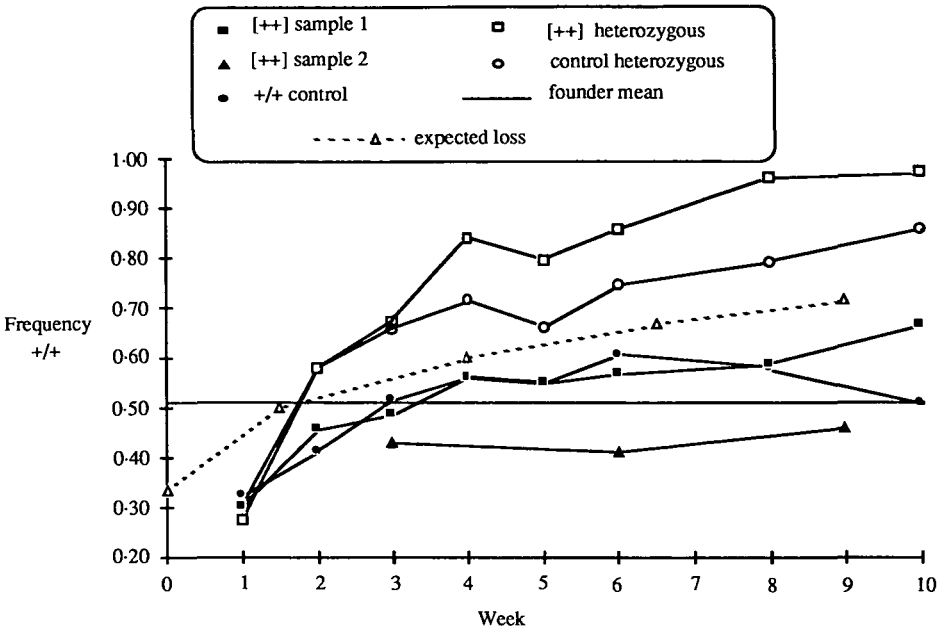

Fig. 4. Frequency of the +/+ genotype in [B+] cages from two-year sample in experiment 2.

isolation. In this case we extracted only 12 chromosomes from the [++] cages. However, four of these cages were set up in duplicate. In addition, we set up four cages containing chromosomes extracted from bottle populations of a wild type Hunter Valley population to act as a control. An extra four cages containing mixtures of chromosomes from the Hunter Valley populations and four from [++] populations were also set up. We sampled weekly initially, and then every two weeks, to monitor more closely the cage frequencies.

None of the extracted chromosomes appeared to increase in frequency beyond the expectations from the founding chromosomes. In Fig. 4 we have averaged over the different classes to show the frequencies of the +/+ genotype over the course of 10 weeks of sampling. The figure also shows the expected rise in frequency of this genotype due just to lethality of the *Cy* homozygote, assuming a generation interval of 2·5 weeks after an initial generation of 1·5 weeks.

The results show an immediate differentiation between the single-chromosome population (filled symbols) and multiple-chromosome populations (unfilled). This result is as commonly found in [B+] populations (Sved & Ayala, 1970; Sperlich & Karlich, 1970). On the other hand there is little difference between the chromosomes extracted from the dichromosomal population (squares) and those extracted from the wild-type populations (circles), except perhaps in the final sample. Unfortunately the counts for the final sample were made after the populations had been discarded, so that it was not possible to see whether the differences indicate any substantial equilibrium differences. In view of the close agreement in all other samples it seems likely that this sample represents a chance fluctuation. Averaging over all samples, there is little difference between the two classes of chromosomes. There appears to be a significant difference between the two classes of mixed chromosomes, which may be attributable to inbreeding in the Hunter Valley bottle population.

In view of the comparatively small number (12) of chromosomes involved in the second extraction, we isolated an additional set of eight chromosomes. The [B+] cages in this set were sampled less frequently than in the previous set. Once again no high fitness chromosomes were found. Averages are shown in Fig. 4 (closed triangles). The average fitness for the founder chromosomes is also shown, and the homozygous samples do not differ systematically from this.

4. Analysis

The results given in the previous section have shown little, if any, increase in the fitness of wild type chromosomes extracted from the dichromosomal populations. Thus our primary conclusion is that deleterious genes of *large* effect are probably a minor cause of the heterosis found in [B+] populations. Our objective in this analysis is to determine how small the selective values need to be for our results to be compatible with the dominance model.

In the accompanying paper (Sved & Wilton, 1989), various combinations of gene number, selective value and degree of dominance have been evaluated to determine which can account for chromosomal heterosis and at the same time be capable of being

maintained by mutation. Our purpose in the present section is to calculate the consequences of these numerical models for the dichromosomal experiment. We do this explicitly for the case of no linkage, but need to turn to computer simulation to take linkage into account.

We use the usual single locus notation for a deleterious gene d (Sved & Wilton, 1989, Table 1). We take the values of s, the selection coefficient, and h, the degree of dominance, as fixed. In terms of these values we then calculate the expected rate of loss of deleterious genes from the [+ +] population.

In setting up the [+ +] population, two chromosomes are drawn from a natural population. Let us denote these as having respectively i and j deleterious genes. The first of these chromosomes has selective value in homozygous condition

$$w_{\text{hom}} = \frac{1}{B}\left[\frac{1-s}{1-hs}\right]^i,$$

where B is a constant selective value associated with the balancer chromosome used for the BE test [Sved & Wilton (1989), equation (1)]. The second chromosome has an equivalent selective value with j replacing i. Assuming that i and j are binomial variates, the expected values of each of these expressions is

$$\bar{w}_{\text{hom}} = \frac{1}{B}\left[1 - sq\left(\frac{1-h}{1-hs}\right)\right]^L, \tag{1}$$

where q is the mean frequency of deleterious gene at each locus in the population, and L is the total number of loci [Sved & Wilton (1989), equation (2)].

These two chromosomes are introduced into a [+ +] population. The frequency of all $i+j$ deleterious genes contributed by the two chromosomes will be 0·5 initially. Sampling from the population at this stage would give a mean number of 0·5 $(i+j)$ deleterious genes per chromosome.

After a period of selection, the frequency of deleterious genes in the [+ +] population will be reduced to a frequency which we will denote as Q. Note that Q is the frequency of deleterious genes in the [+ +] population, initially 0·5, and is unrelated to q, the frequency of deleterious genes in the population from which the [+ +] population was founded. The value of Q will in general be a function of s, h, the degree of linkage, and the number of generations.

We must now consider the second sampling process, the sampling of chromosomes from the [+ +] population. The measurement of selective value is made by setting up [B+] populations from single chromosomes sampled from the [+ +] population. As previously mentioned, we assume linkage equilibrium initially. Sampling from the [+ +] population will yield a binomial distribution, with the probability of sampling x deleterious genes from amongst $i+j$ equal to

$$f(x) = C_x^{i+j} Q^x (1-Q)^{i+j+x}.$$

The selective value of a chromosome homozygote with x deleterious genes is

$$w_x = \frac{1}{B}\left[\frac{1-s}{1-hs}\right]^x.$$

The mean selective value of the sample becomes

$$\bar{w}_{ij} = \sum_{x=0}^{i+j} f(x)\, w_x,$$

which reduces to

$$\bar{w}_{ij} = \frac{1}{B}\left[1 - sQ\left(\frac{1-h}{1-hs}\right)\right]^{i+j}.$$

The mean selective value at the stage at which the [+ +] population is founded can be derived by putting $Q = 0{\cdot}5$.

Summing over all possible values of i and j gives the overall expected value of chromosomes isolated from all [+ +] populations as

$$\bar{w}_{[++]} = \sum_{i=0}^{L}\sum_{j=0}^{L} \frac{1}{B}\left[1 - sQ\left(\frac{1-h}{1-hs}\right)\right]^{i+j} \times C_i^L q^i (1-q)^{L-i} C_j^L q^j (1-q)^{L-j},$$

and this simplifies to

$$\bar{w}_{[++]} = \frac{1}{B}\left[1 - sqQ\left(\frac{1-h}{1-hs}\right)\right]^{2L}. \tag{2}$$

Note that when $Q = 0{\cdot}5$, at the founding of the [+ +] population, this becomes

$$\bar{w}_{[++]}(t=0) = \frac{1}{B}\left[1 - sq\left(\frac{1-h}{2(1-hs)}\right)\right]^{2L}. \tag{3}$$

The right-hand side of (3) is similar, although not identical, to equation (1). The process of sampling from a [+ +] population yields a distribution which differs slightly from that used to set up the population. If selective values were additive rather than multiplicative, there would be no difference between the two mean values.

The principal quantity of interest in this discussion is the rise in fitness as deleterious genes are eliminated. This can be expressed in two different ways. The first is as the ratio of $\bar{w}_{[++]}$ given by (2) relative to $\bar{w}_{[++]}$ $(t=0)$ given by (3). This is the rise in fitness of chromosomes sampled from the [+ +] cage over the course of the experiment. However there is little difference if the denominator in this expression is, instead, $\bar{w}_{\text{hom}}$ as given by (1). This compares the mean fitness of chromosomes sampled from the [+ +] population with the mean fitness of chromosomes used to found the population. We use the latter in the present case, giving

$$w = \left[1 - sqQ\left(\frac{1-h}{1-hs}\right)\right]^{2L} \Big/ \left[1 - sq\left(\frac{1-h}{1-hs}\right)\right]^{L}. \tag{4}$$

Note that this ratio eliminates the effect of the balancer chromosome (B).

https://doi.org/10.1017/S0016672300028019 Published online by Cambridge University Press

Table 2. *Expected mean fitness of wild-type chromosome homozygotes extracted from a [+ +] population after 32 generations, assuming a starting mean fitness of 0·2, multiplicative interaction, and no linkage*

	s						
h	1/2	1/4	1/8	1/16	1/32	1/64	1/128
1/2·5	0·9985	0·9293	0·6616	0·4148	0·2954	0·2441	0·2211
1/4	0·9852	0·8546	0·6001	0·3964	0·2913	0·2432	0·2209
1/8	0·9351	0·7716	0·5534	0·3825	0·2880	0·2424	0·2207
1/16	0·8860	0·7270	0·5319	0·3760	0·2865	0·2420	0·2206
1/32	0·8559	0·7048	0·5216	0·3729	0·2857	0·2419	0·2206
1/64	0·8397	0·6937	0·5166	0·3714	0·2853	0·2418	0·2206
1/128	0·8315	0·6883	0·5142	0·3706	0·2851	0·2417	0·2205
0	0·8231	0·6828	0·5117	0·3698	0·2849	0·2417	0·2205

(i) *Calculations*

There are altogether 5 parameters involved in equation (4) which predicts the course of events in the [+ +] population. We now fix the values of s and h, or rather we consider a range of values of s and h. Fixing the value of L then allows us to calculate the value of q from equation (1), assuming $B = 1$ (see below). The value of Q, the frequency of deleterious genes after selection in the [+ +] population, is determined by s and h and by the number of generations elapsed in the [+ +] population. This generation number is taken to be 32, which for experiment 2 is a conservative underestimate. No simple formula can be given for Q, but it is easy to calculate by iteration the expected frequency, provided it is assumed that there are no linkage effects. The various parameters are then substituted into equation (4) to give the expected fitness.

We will make a simplifying assumption in the present calculations compared to the treatment given by Sved & Wilton (1989, equations (3), (4) and table 2), in which fitnesses from heterozygous cages entered directly into the calculations. In the dichromosomal experiments of the present paper, heterozygous cages were not used. All measurements of fitness were made in [B+] cages, relative to the *Cy* balancer chromosome. We alluded in the Methods and Materials section to the complication introduced by the fitness of this chromosome. Since it is not present in the [+ +] populations, its fitness is not relevant to the rate of loss of deleterious genes. The fitness of the *Cy* chromosome, B, enters into the argument only because all measurements are made relative to the *Cy*/+ genotype. Yet the quantity B cancels out from equation (4) which predicts the rise in fitness of chromosome homozygotes. Thus the same relative rise in fitness should be found regardless of whether fitness is measured relative to a low fitness or a high fitness genotype. The simplification we have therefore made is to put $B = 1$ in equation (1), thereby allowing the parameter q to be estimated directly from this equation. Equation (4) then reduces to

$$w = \left[1 - sqQ\left(\frac{1-h}{1-hs}\right)\right]^{2L} \Big/ (0{\cdot}2). \qquad (5)$$

The numerator of this expression is $\bar{w}_{[++]}$ from equation (2). Calculations can equivalently be made in terms of either of the parameters w or $\bar{w}_{[++]}$.

Numerical results are shown in Table 2, which gives the value of $\bar{w}_{[++]}$ as calculated from equation (2), which is equivalent to the numerator of equations (4) or (5). There is a high dependence on the value of s, and a low dependence on h excepting in the upper range of s and h values. The change in fitness is approximately proportional to s over most of the range.

An important conclusion from the paper of Sved & Wilton (1989) concerns the role of partial dominance. Deleterious genes showing partial dominance are expected to occur at a much lower equilibrium than fully recessive deleterious genes of equivalent severity, thereby making a lower contribution to heterosis. For the [+ +] population, however, the role played by partial dominance is a minor one. Given that a deleterious gene exists in the [+ +] population, it is clear from Table 2 that it makes little difference to its expected rate of loss whether it is fully recessive or not.

If we interpret the results of this paper as showing that the selective value in [+ +] populations has not increased by a factor of 50%, then the maximum value which s could take is around 1/32. However it must be remembered that this calculation does not take linkage into account. We attempt to do this in the following section by computer simulation.

(ii) *Computer simulation*

The primary aim of the simulations reported below is to estimate the factor by which linkage retards the

https://doi.org/10.1017/S0016672300028019 Published online by Cambridge University Press

Table 3. *Mean homozygous fitness and standard deviations computed for unlinked loci, assuming $h = 1/16$*

N	s... 1/4 Lq... 6·8	1/16 27·4	1/64 109·8
16	0·7733 ± 0·0024	0·3294 ± 0·0018	0·2281 ± 0·0008
64	0·7485 ± 0·0020	0·3570 ± 0·0016	0·2378 ± 0·0008
256	0·7303 ± 0·0020	0·3739 ± 0·0016	—
Expected	0·7270	0·3760	0·2420

fixation of deleterious genes in the [+ +] population. The simulation involves only the [+ +] population. The amount of recombination does not affect the outcome in the [B+] populations, since the balancer chromosome is assumed to suppress all recombination. Therefore we assume that the outcome in these populations is strictly determined by the chromosome fitness, which can be calculated directly under the assumption of multiplicative selective values.

A secondary aim of the computations is to test for possible effects of finite population size. The theory leading to the expectations of Table 2 is based on an infinite model, which cannot easily be extended to take into account finite size effects. By comparison, a Monte-Carlo simulation can only be carried out for a finite size population. However by increasing the population size in the simulation it should be possible to test whether the selective value is becoming asymptotically equal to the infinite-size value. Such equality serves as a test of both the calculations and of the computer program.

The initial simulations were carried out with free recombination between all loci. Three different populations sizes were used, $N = 16$, 64 and 256. In view of the low dependence of fitness change on the degree of dominance h, which we confirmed for simulations involving linkage, we show only the results assuming $h = 1/16$. Most of the range of selective values was spanned using $s = 1/4$, $1/16$ and $1/64$. For each value of s, we used equation (1) to estimate the mean number of loci per chromosome consistent with an average homozygous fitness of 0·2, giving Lq values of 6·8, 27·4 and 109·8 respectively. We then generated randomly two chromosomes for the [+ +] population, using binomial sampling from the above mean numbers to determine the actual number of loci per chromosome. Each such population was then run four times for 32 generations, and the mean selective value of chromosome homozygotes calculated at the end of the run. Altogether 1024 such populations were generated for each s and N value.

Mean fitness values and their standard errors are given in Table 3. The expected selective value from Table 2 is also given. There is a small but noticeable effect of population size, whose direction depends on the value of s. The values for the highest population sizes are in good agreement with their calculated expectation. The one missing value in the table, corresponding to the highest N and Lq values, would have taken several days to compute, and would be expected to provide little additional information.

Standard deviations from individual runs are not shown in the table. However the averages in Table 3 are based on 1024 values (each of which is in turn the average from four runs), so that the standard deviation of each set of four can be obtained by multiplying the standard error by $\sqrt{1024} = 32$. It can be seen that the variability between runs is high. The main reason for this high variability is the sampling of genes when [+ +] populations are generated. Setting up the populations requires sampling of chromosomes from a large pool. When a small number of genes is being sampled, i.e. low value of Lq in Table 3, there is considerable variability in the number of deleterious genes sampled in different populations. This variability is not a problem in the dichromosomal experiments provided that the same [+ +] populations are measured at the start and finish of the experiment. Note that the choice of Lq is made taking this variability into account (equation 2).

Table 4. *Mean homozygous fitness computed for the model of one crossover per chromosome arm. Values shown in parentheses are the percentages of retardation compared to the model of free recombination (see text)*

N	s... 1/4 Lq... 6·8	1/16 27·4	1/64 109·8
16	0·4914 (50·8)	0·2672 (51·9)	0·2159 (56·5)
64	0·5325 (60·6)	0·2783 (49·9)	0·2201 (53·2)
256	0·5683 (69·5)	0·2855 (49·2)	—

https://doi.org/10.1017/S0016672300028019 Published online by Cambridge University Press

Simulations involving linkage are shown in Table 4. The simulations are based on the same number of replications as those involving unlinked loci. Table 4 shows the mean fitness obtained for each combination of N and s values, together with a measurement of the proportionate increase under linkage. For instance, for the measurement corresponding to $N = 16$, $s = 1/4$, from Table 3 the increase in mean fitness with free recombination is $0{\cdot}7733 - 0{\cdot}2 = 0{\cdot}5723$. The corresponding increase with linkage from Table 4 is $0{\cdot}4914 - 0{\cdot}2 = 0{\cdot}2914$. This represents a proportion $0{\cdot}2914/0{\cdot}5723 = 0{\cdot}508$ of the increase under free recombination, which is tabled in parentheses in Table 4.

The increase in mean fitness with linkage is approximately 50–60% of the amount with free recombination, showing that linkage has a substantial retarding effect on the elimination of linked genes. Surprisingly, the effect is not very different for small numbers and large numbers of genes. The mean change in gene frequency is smaller when large numbers of genes are involved, which appears to compensate for any increase in linkage effects. Furthermore, population size has an unpredictable effect on the outcome, with increase in population size sometimes showing increased linkage effects and sometimes reduced effects.

5. Discussion

Table 4 enables us to make a judgement about how low a value of s can be ruled out from our experiments. Clearly the predicted increase with $s = 1/4$ (column 1) would have been detected in our experiments. On the other hand, we probably cannot rule out a value of $s = 1/64$, since we could not expect to detect increases of 10% or less. The median value of $s = 1/16$ is more equivocal, although it seems likely that we would have been able to detect a rise of this magnitude (30–40%). As a conservative solution, we would be confident of being able to rule out values of s greater than $1/16$, and perhaps even values of half of this amount.

In order to see the implications of this result for mutation-selection balance models, we need to turn to the results established by Sved & Wilton (1988). There are clearly no problems of reconciling a value $s = 1/32$ with a mutation rate of less than 10^{-5} provided that the value of h is low [Sved & Wilton (1989), tables 3, 4]. However the results argue against agreement with models having substantial values of hs such as estimated for the viability mutants studied by Crow (1979). The possibility has been raised by Sved & Wilton (1989) that these viability mutants may have reasonably high values of s, precisely the class which the dichromosomal experiment has ruled out as a general explanation for chromosomal heterosis. Thus relatively low levels of h, e.g. $h \leqslant 1/8$, could explain both the results of Crow and those of the present experiment.

In interpreting the results from the dichromosomal experiment, one possibility which must be taken into account is that new deleterious mutations may be occurring to mask the loss of deleterious genotypes in the [+ +] populations. Theoretically this effect should not be a large one. Natural populations may be assumed to be normally in equilibrium under a mutation–selection balance. Such an equilibrium is not expected to apply in [+ +] populations. This is because the selection component is much higher in [+ +] populations, owing to the artificially high frequency of any deleterious genes. New mutations could mask some, but presumably not the majority, of the expected loss of deleterious genes.

The possibility also exists of dysgenic effects increasing the mutation rates compared to natural populations, although the [+ +] populations were set up in a way to minimise dysgenic effects (see Materials and Methods). The rate of occurrence of lethal mutations, which was studied in detail in the first experiment, provides some evidence on this point. From 211 chromosomes sampled from the [+ +] cages, 23 independent new lethal mutations were found. Based on the expected level of 0·005 lethal mutations per generation (Simmons & Crow, 1977) and the average number of around 20 generations estimated to have elapsed before sampling, this figure appears not to be significantly above expectation. Thus lethals provide no evidence for any dysgenic effects.

Sved & Wilton (1989) have emphasized that the calculations on numbers and intensities of deleterious genotypes are independent of whether the dominance or overdominance model is assumed. However the predictions from the dichromosomal experiment are quite different for the two models. Under the overdominance model, the loss of deleterious genotypes is not expected to follow the same dynamics as predicted by the dominance model. Therefore the results of this paper do not rule out the possibility that there are relatively small numbers of overdominant loci of large effect which are responsible for the bulk of the inbreeding depression.

The argument against genes of large effect for the dominance model means that there must instead be a large number of genes affecting fitness per chromosome. Table 2 of Sved & Wilton (1989) indicates a minimum of 50 or so deleterious genes per chromosome. Estimates of the extent of polymorphism at the molecular level (Kreitman, 1983) easily leave room for such a number. The challenge of elucidating at the single gene level which mutations are the deleterious ones is a formidable one.

Comments from several members of the British Population Genetics Group are acknowledged following an oral presentation of an earlier version of this paper. The work was supported by a grant from the Australian Research Grants Scheme.

References

Collins, G. N. (1921). Dominance and the vigor of first generation hybrids. *American Naturalist* **55**, 116–133.

Crow, J. F. (1979). Minor viability mutants in *Drosophila*. *Genetics* **92**, s165–s172.

Crow, J. F. & Simmons, M. J. (1983). The mutation load in *Drosophila*. In *The Genetics and Biology of Drosophila*, vol. III (ed. M. Ashburner and J. N. Thompson). New York: Academic Press.

Davenport, C. B. (1908). Degeneration, albinism and inbreeding. *Science* **28**, 454–455.

Dobzhansky, Th. & Spassky, B. (1954). Genetics of natural populations. XXII. A comparison of the concealed variability in *Drosophila prosaltans* with that in other species. *Genetics* **39**, 472–487.

East, E. M. (1908). Inbreeding Corn. Report of the Connecticut Agricultural Experimental Station for 1907, pp. 419–428.

East, E. M. & Hayes, H. K. (1912). Heterozygosis in evolution in plant breeding. U.S. Department of Agriculture, *Bureau of Plant Industry Bulletin* **243**, 58.

Fristrom, J. W. & Clegg, M. T. (1988). *Principles of Genetics*, 2nd edn. New York: Chiron.

Gardner, C. O. (1963). Estimates of genetic parameters in cross-fertilizing plants and their implications in plant breeding. In *Statistical Genetics and Plant Breeding* (ed. W. D. Hanson and H. F. Robinson). Washington: National Academy of Sciences.

Gowen, J. W. (1952). Hybrid vigor in *Drosophila*, pp. 474–493. In *Heterosis* (ed. J. W. Gowen). New York: Hafner

Hull, F. H. (1952). Recurrent selection and overdominance, pp. 451–473. In *Heterosis* (ed. J. W. Gowen). New York: Hafner.

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

Jones, D. F. (1917). Dominance of linked factors as a means of accounting for heterosis. *Genetics* **2**, 466–479.

Kidwell, M. G., Kidwell, J. F. & Sved, J. A. (1977). Hybrid dysgenesis in *Drosophila melanogaster*: a syndrome of aberrant traits including mutations, sterility and male recombination. *Genetics* **86**, 813–833.

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge University Press.

Kreitman, M. (1983). Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*. *Nature* **304**, 412–417.

Lewontin, R. C. (1974). *The Genetic Basis of Evolutionary Change*. New York: Columbia University Press.

Pascoe, L. (1985). Fitness components of chromosomal homozygotes in *Drosophila melanogaster* in age-structured populations. PhD Thesis, The University of Sydney.

Prout, T. (1965). The evolution of fitnesses from genotypic frequencies. *Evolution* **19**, 546–551.

Schull, G. H. (1908). The composition of a field of maize. *Report of the American Breeders Association* **5**, 51–59.

Schull, G. H. (1911). The genotypes of maize. *American Naturalist* **45**, 234–252.

Simmons, M. J. & Crow, J. F. (1977). Mutations affecting fitness in *Drosophila* populations. *Annual Review of Genetics* **12**, 289–328.

Sperlich, D. & Karlik, A. (1970). The genetic conditions in heterozygous and homozygous populations of *Drosophila*. I. The fate of alien chromosomes. *Genetica* **41**, 265–304.

Sprague, G. F., Russell, W. A. & Penny, L. H. (1959). Further studies on convergent evolution in corn. *Genetics* **44**, 341–346.

Sved, J. A. (1971). An estimate of heterosis in *Drosophila melanogaster*. *Genetical Research* **18**, 97–105.

Sved, J. A. & Ayala, F. J. (1970). A population cage test for heterosis in *Drosophila pseudoobscura*. *Genetics* **66**, 97–113.

Sved, J. A. & Wilton, A. N. (1989). Inbreeding depression and the maintenance of deleterious genes by mutation: model of a *Drosophila* chromosome. *Genetical Research* **53**, 119–128.

Wilton, A. N. (1980). Studies of the nature of chromosomal heterosis in *Drosophila melanogaster*. PhD Thesis, University of Sydney.

Yamazaki, T. & Maruyama, T. (1971). Evidence for the neutral hypothesis of protein polymorphism. *Science* **178**, 56–58.