FITNESS IN DROSOPHILA

Contents

0.1. **Preface.**

This section of the PIFFLE describes experiments I have been involved in using Drosophila to measure selection. This field of what might be described as laboratory population genetics was fairly mainstream in the 1950s - 1980s. It gradually declined in importance as molecular biology took over and made it possible to study populations in a wide range of species. Perhaps the field will rise again, as I believe that many problems of how natural selection works cannot be studied without an experimental rather than observational approach.

1. INBREEDING AND EGG-TO-ADULT SURVIVAL

In the section on genetic load, I have written about the predicted inbreeding depression under a heterozygote advantage model. Under this model, homozygotes are always at a selective disadvantage to heterozygotes. Obviously if many loci with homozygous disadvantage are made homozygous at the same time, as happens with inbreeding, this should have selective consequences. This seems a simple enough prediction, but somehow it was one that was lost sight of during the debates on genetic load.

Having decided that inbreeding provided the critical test for heterozygote advantage, I decided that I should be involved in experimental investigation of inbreeding. The obvious organism to use in such tests was Drosophila. I don't remember much of the reasoning behind this. I had no experience with Drosophila, beyond failing in an undergraduate exercise where we were asked to map a mutant gene. There were no Drosophila labs at Stanford, where I was working at the time. Fortunately David Perkins offered me space in his Neurospora lab, where work on Drosophila behaviour was being carried out by Seymour Kessler, a postdoc in the Biology Department at the time.

Somehow I must have eventually gotten in touch with Theodosius Dobzhansky, the leading Drosophila population geneticist of his time. I was in Dobzhansky's good books for coming down on his side in the genetic load debates, and he offered me a postdoctoral fellowship in his lab at the Rockefeller University to do these experiments. I'm sure he didn't realise what a lousy experimentalist he was getting.

1.1. **Balancer chromosomes.**

I should start by describing briefly the procedure used to measure inbreeding in Drosophila. H.J. Muller, one of the pioneers of Drosophila genetics, had invented the so-called balancer chromosome. This is a chromosome marked with a dominant mutation, containing one or more inversions to suppress crossingover, and typically lethal in homozygous condition.

The theory behind the use of this chromosome is very simple. I'll call this chromosome Cy. This is actually the designation of a particular (second) chromosome balancer chromosome in *Drosophila melanogaster*, or more correctly one of a number of such balancer chromosomes. The Cy gene leads to Curly wings. One can then refer to a genotype that contains one copy of this balancer chromosome and one normal chromosome as

$Cy\ /\ +$

This is exactly the nomenclature for a heterozygous genotype at a single locus, but now it refers to an entire chromosome.

1.2. **Measuring viability.**

Muller used the balancer chromosome primarily to study irradiation - he was the discoverer of the mutagenic effects of irradiation. But it has numerous other uses, one of which is to produce controlled levels of inbreeding.

Suppose one goes into a population of Drosophila and collects some male flies. These are crossed to females from a stock containing a balancer chromosome. This cross can be shown as:

Cy/l x $+/+$

The genotype containing the balancer stock is denoted as Cy/l, signifying that it contains a second lethal gene somewhere on the chromosome. It's use is simply to maintain the balancer chromosome since a genotype such as Cy/Cy cannot survive.

The offspring from this cross are expected to be either $Cy/+$ or $l/+$. If one chooses a single $Cy/+$ offspring and crosses it again to the balancer stock, the cross can be designated:

Cy/l x $Cy/+$ (single parent)

Any offspring containing the balancer chromosome must have the genotype $Cy/+$, since Cy/Cy flies don't survive. Furthermore all such $Cy/+$ flies must contain exactly the same wild type + chromosome, produced by DNA replication from the original chromosome at the one-cell stage of the single $Cy/+$ parent. Now one intercrosses female and male $Cy/+$ offspring:

$Cy/+$ x $Cy/+$

to give

$Cy/+$ and $+/+$

These $+/+$ flies are remarkable in that they must be homozygous for every wild type gene on the entire chromosome. Furthermore the expected ratio of offspring is 2 $Cy/+$: 1 $+/+$. Significantly less than this fraction of $+/+$ indicates a disadvantage of one or more homozygous genes.

The technique as outlined is actually an inbreeding technique. All individuals in the $Cy/+$ x $Cy/+$ cross are at least half sibs, sharing one, but possibly two, parents. However conventional half-sib mating just gives a probability $\frac{1}{8}$ of homozygosity without specifying where in the genome it occurs. The balancer technique ensures homozygosity of a single whole chromosome. There is also some inbreeding on other chromosomes.

Furthermore the ratio of emerging offspring allows direct measurement of the *viability*, or egg-to-adult survival, of chromosome homozygotes. Note that the viability is measured against a standard genotype containing the balancer chromosome. One can do a control for possible effects of the balancer chromosome by crossing $Cy/+$ flies containing different wild type chromosomes, thereby measuring the viability of $Cy/+$ individuals against normal $+/+$ heterozygotes.

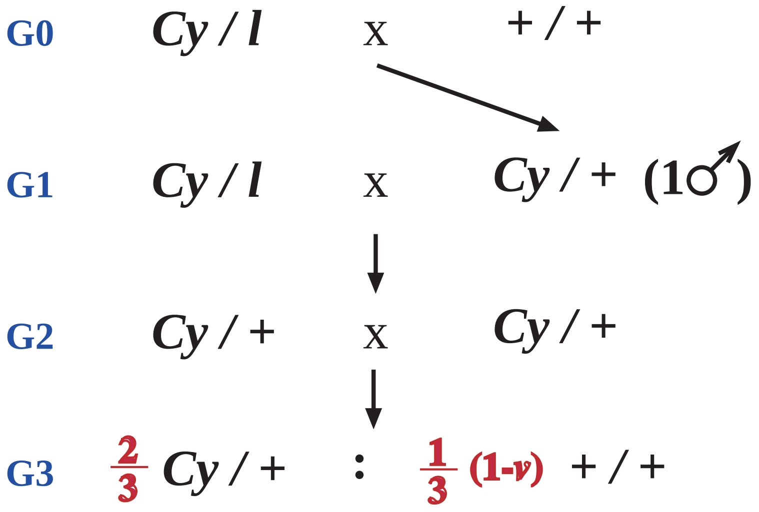

FIGURE 1. The crossing procedure for measuring viabilities of whole chromosomes

It's convenient to summarise the whole procedure as in Figure 1. Generation designations, in blue, are arbitrarily shown as starting from zero. The survival probability of a $+/+$ genotype compared to $Cy/+$ is designated as $1 - v$, where v is a 'selection coefficient'. Note the rather strange frequencies $\frac{2}{3}$ and $\frac{1}{3}$. Normally the offspring produced by a cross such as that shown in G2 would be in the ratio $\frac{1}{4} : \frac{2}{4} : \frac{1}{4}$, but because the Cy/Cy offspring don't survive, the frequencies are normalised with a 3 rather than 4.

1.3. **Some results.**

Procedures such as those shown in Figure 1 have been used for a number of species of Drosophila. These experiments were typically done in the 1960s and 1970s, and a convenient review was given by Simmons and Crow (1977) [9]. The first notable feature is the large number of recessive lethal mutations uncovered. Between 20% and 30% of crosses such as shown in Figure 1, possibly more for the larger chromosomes, give no wild type progeny. These presumably reflect mutations in vital genes that show little effect in single dose but are lethal in homozygous condition.

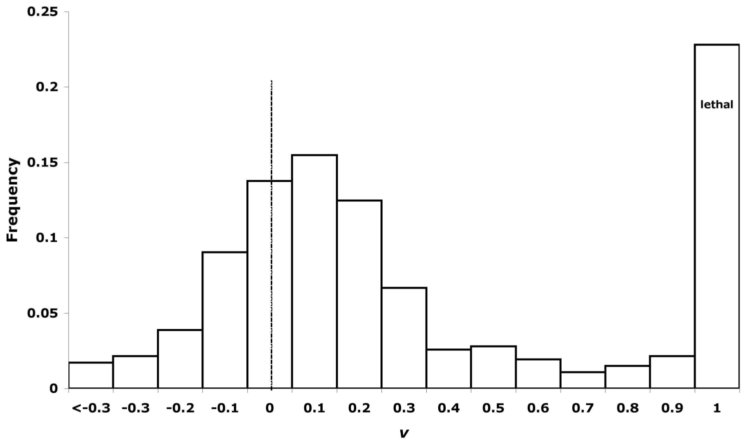

Figure 2. The distribution of selection coefficients against chromosome homozygotes

The overall distribution of v values from some studies of chromosome 2 in *D. melanogaster* is shown in Figure 2. The dotted line marks the point where there is no selection. There is clearly a substantial bias to the right of this point, including the right-most class, which are the lethal chromosomes.

The peak of non-lethal chromosomes lies to the right of the zero point, at around $v = 0.1$. The mean value of v is around 0.14 for non-lethal chromosomes, and 0.33 for all chromosomes. The negative values of v in Figure 2 may represent 'superfit' chromosome homozygotes, but are more likely due to chance. The v values are obtained relative to *Cy* balancer chromosome heterozygotes. However the control crosses show that such genotypes have no significant reduction in viability compared to the normal wild type chromosomes.

In summary, therefore, there is some reduction in viability of chromosome homozygotes, ie a limited amount of inbreeding depression.

Alternatively, looking at the situation from the point of view of chromosomal heterozygotes, such inbreeding depression may be described as chromosomal heterosis. However it seems not nearly enough to be compatible with heterozygote advantage at thousands of loci along the chromosome.

1.4. **Competitive experiments.**

Despite these apparently negative results from earlier studies, I was keen to repeat some of these experiments. The reason was the various 'soft selection' models put forward by Wallace [19], King [3], Milkman [7] and our group [15], as described in the chapter on genetic loads. In these models selection is attributable to density-dependent competition for limited resources, with higher overall death rates leading to higher levels of selection at individual loci. I was therefore convinced that much higher levels of inbreeding depression would be revealed if we increased the crowding in these experiments.

I'm not able to give any results from these experiments, except to say that they never revealed any great effects of crowding on selective values. Nothing was ever published, for which there are several reasons. Chief of these was probably the negative nature of the results. A second reason was a lack of ability on my part to follow through the experiments to publication. In my defense, however, these experiments were overtaken by a new and more promising set of experiments that are detailed in the next section.

Although nothing was published, I'm fairly sure that enough was done to show that there is a genuinely small effect of crowding on the selective values. I found this first in the rather primitive experiments that I did using *D. melanogaster* in Stanford, just using larger numbers of parents per crossing tube to increase levels of crowding.

When I got to Rockefeller University, Dobzhansky showed me how to do these experiments properly. The G2 crosses of Figure 1 were used to produce large numbers of eggs. These eggs were counted, and placed in different numbers, from memory something like 20 per tube for the low density experiments, 100 per tube, 500 per tube, etc. Each tube held a measured amount of food, enough for 50-100 adults to emerge. There was no doubt, therefore, that the amount of food was a limiting resource.

I have forgotten how many different chromosome were studied in this way. However I believe that at least some chromosomes were used which gave reduced viabilities at low density, and that these viabilities were unchanged at the higher densities.

I should mention that these experiments were done using *D. pseudoobscura*. Dobzhansky was somewhat contemptuous of *D. melanogaster*, calling it a 'garbage species'. There is no doubt that *D. melanogaster* is commensurate with humans, and is therefore an 'unnatural' species. However I believe that some time after we did our experiments evidence emerged that *D. pseudoobscura* is also not quite as 'pure' as Dobzhansky liked to believe, with higher numbers being trapped in areas with more garbage. I was also surprised when I first saw *D. pseudoobscura* in the wild, how large they were compared to the scrawny specimens that we grew in the lab.

In summary, the competitive experiments on egg-to-adult viability showed little, if any, effect on ratios of homozygous flies emerging. Clearly there was competition for resources, but who knows what goes on in the tubes while larvae emerge, feed, pupate, and then eclose as adults. Maybe the first to emerge get all the food. Anyway I am still intrigued at the negative nature of the results. If homozygous flies have a reduced probability of surviving, and this is not related to density, what causes the inviability? I'll outline an experiment in the last section of this chapter that was designed to investigate this point.

I should also mention here a result from Leigh Pascoe, who was a PhD student I supervised in the 1980s. Leigh's suprising result was that crowding not of the offspring but of the parents led to a substantial viability effect. I'll come back to this result in the section of fitness components later in this chapter.

2. THE BALANCER EQUILIBRATION TECHNIQUE

Although the crosses of Figure 1 show only three generations, in practice it was often convenient to throw in a few extra generations, partly to build up numbers and partly to time the experiments better. So the actual crossing program could be described as in Figure 3.

For the generation labelled G2a, the actual cross used to estimate viabilities, it was important to ensure that the females in the cross were virgin. This is done in Drosophila by collecting young females, usually less than 6 hours old, to be sure that they have not mated with any old male that they first come into contact with. Otherwise the cross that you think you are setting up may not be the actual cross.

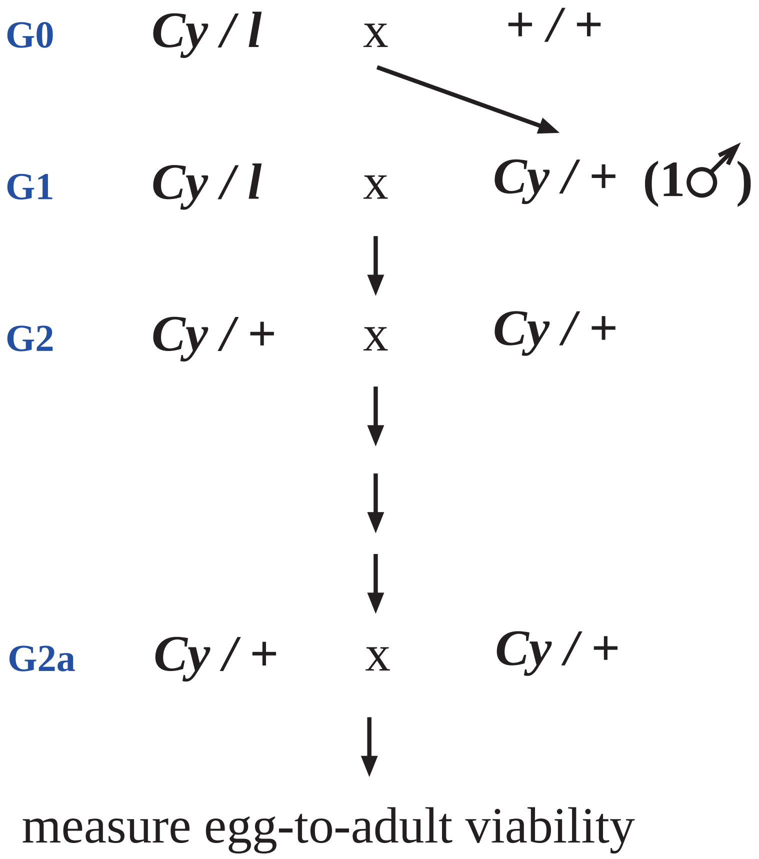

FIGURE 3. Extended crossing program

For those intermediate generations, the ones just marked with arrows, all that was done was to collect old $Cy/+$ flies and put them into new bottles. Under these circumstances it is possible that the $Cy/+$ females will in fact have mated with $+/+$ males. Sometimes to save work the flies were simply tipped over, allowing $+/+$ flies to contribute.

What gradually became clear was that $+/+$ flies didn't seem to be contributing as much as might have been expected to the next generations. Since Cy/Cy is lethal, one might have expected that the frequency of the Cy balancer chromosome would gradually drop over the generations. That is, unless this was opposed by a corresponding selection against the $+/+$ genotypes.

At this stage the penny dropped. Here we had a great way of measuring the fitness of $+/+$ flies. The theory is very simple. If $+/+$ is at a disadvantage to $Cy/+$, then the 'selective value' of the $+/+$ genotype compared to $Cy/+$ may be written as 1 - s, where s is the so-called 'selection coefficient'. The selective values, or fitnesses, of the three genotypes may be written as:

Genotype:	Cy/Cy	Cy / +	+ / +
Fitness:	0	1	1 - s

Anyone familiar with elementary population genetics theory will recognise this as a classical heterozygous advantage model, with the heterozygote at an advantage to both homozygotes. The Cy chromosome, despite its lethality, is expected to come to an equilibrium, at frequency 1 / (1 + s). The observed equilibrium frequencies may thereby be used to estimate s. This is the basis for the Balancer Equilibration (BE) technique.

2.1. **Population cage experiments.**

The experiment to test balancer equilibration required the setting up of 'populations' rather than just 'crosses'. We could have just tipped flies over from old bottles to new bottles every couple of weeks, more or less as was done in Figure 3. However a better way of doing it is to use 'population cages'. Francisco Ayala, who was working in Dobzhansky's lab at the time, had designed a very simple and suitable cage, based on a plastic lunch box, that enabled large numbers of populations to be carried along at the same time. An artists's interpretation of such a cage is shown in Figure 4. Six food bottles are attached, and are exchanged with new food in rotation over a 3-week period. The body of the cage is often unchanged over the life of the experiment, a few months in this case.

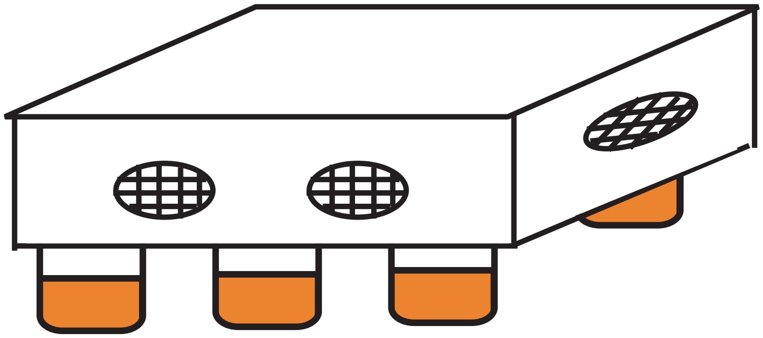

Figure 4. A Drosophila cage

Francisco agreed to collaborate on this experiment [14]. We set up 36 cages, two duplicates of each of 18 founder chromosomes. Lethal chromosomes were excluded, since they would show nothing extra in this type of experiment. Most of the 18 chromosomes gave high viability estimates.

Despite the high viability estimates, in no case was a balancer chromosome eliminated from any cage, as might have been expected from the lethality of the balancer chromosome homozygotes. Estimates of the fitness of +/+ homozygotes ranged from 50% downwards. The typical chromosome looked something like the following:

Genotype:	Cy/Cy	$Cy/+$	+/+ (homozygote)
Fitness:	0	1	0.4

The supposedly 'superfit' chromosomes of Figure 2 were not really superfit at all.

We also set up 4 cages containing mixtures of wild type chromosomes. The results were totally opposite for these cages. In all cases the balancer chromosome was eliminated, at a rate that indicated that it was selectively disadvantageous in heterozygous condition, in addition to being homozygous lethal. This was the estimate:

Genotype:	Cy/Cy	$Cy/+$	+/+ (heterozygote)
Fitness:	0	0.5	1

In retrospect this was not surprising, given that the inversions are expected to induce fertility problems. But it showed that +/+ homozygotes were at a large selective disadvantage to a genotype that was itself highly unfit.

What is really needed is the fitness of +/+ homozygotes compared to +/+ heterozygotes. Assuming that +/+ homozygotes have a fitness of 40% compared to the $Cy/+$ genotype, which itself has a fitness of 50% compared to +/+ heterozygotes, we estimated this as 40% of 50%, or 20%. This compares to viability estimates of 80-90%. The levels of inbreeding depression uncovered in this experiment are compatible with substantial levels of either heterozygous advantage or homozygous disadvantage at many loci on the chromosome.

The original BE experiment was done in *D. pseudoobscura*. I eventually repeated it for both major autosomes in *D. melanogaster* [11] [13], with some small improvements in the way that frequencies were estimated (see next section). Francisco and colleagues repeated it in other Drosophila species [8] [18]. The experiment was also remarkably similar to one published by Sperlich and Karlik (1970) [10]. BE was an experiment that kept on giving. To the best of my knowledge, no one has ever found a wild type chromosome with anything other than very low homozygous fitness.

2.2. **Sampling the population.**

Originally Francisco and I calculated chromosome frequencies by anaesthetising all flies in the cage and counting adults. There are two disadvantages of this means of censusing the population. The first is the difficulty of the procedure. The second is that the theory for estimating fitness components really requires one to have estimates at the gamete or zygote (one-cell) stage, rather than part way through the life cycle in the adults.

We developed a simpler sampling method as follows. One day before a food jar was changed in the cage, we put in a sampling jar for collection of eggs. Batches of around 100 of these eggs were than spooned into tubes with regular amounts of food, the individuals raised to adulthood and scored for the *Cy* phenotype.

The difficulty of this method is that it still gives frequencies at the adult stage. Although progeny are grown under optimal conditions, there is still expected to be egg-to-adult inviability. However we could correct for this by setting up $Cy/+$ x $Cy/+$ crosses and growing the eggs under the same conditions, thereby giving a correction for the egg-to-adult viability.

As far as we could tell, this sampling method improved the accuracy of the BE technique. In 10-15% of cases, however, we found a rather bizarre result - the frequency of the *Cy* phenotype was significantly higher in the sampling vials than in the viability crosses. This, in theory, means that the $+/+$ parents in the cage are making a negative contribution to the *Cy* counts, leading to a negative estimate for the frequency of $+/+$ individuals in the cage. We didn't really resolve this at the time. This comes up again in the section below entitled 'Components of fitness'.

2.3. **Sex chromosomes.**

The statement concerning low fitness homozygotes should strictly refer to chromosomes other than sex chromosomes, ie. autosomes. But even for sex chromosomes, Alan Wilton, for whom I was PhD supervisor, estimated using the BE technique that wild type chromosome homozygotes, necessarily in the female, were at a 40% disadvantage to wild type heterozygptes [20]. This compares to an 80% disadvantage for autosomes. However the X chromosome is only about half the size of the autosomes in *D. melanogaster*, so that there seems to be rather little difference between X chromosome and autosome disadvantage.

This came as rather a surprise to us. We had assumed that the dynamics of deleterious homozygous genes would be very different on sex chromosomes compared to autosomes. Such deleterious genes on sex chromosomes should be rapidly eliminated since they are present in single dose in the male. We ended up deciding that sex-limited expression, ie. genes that are expressed in the female and not in the male, might be important in this experiment.

2.4. **Components of fitness.**

There is an extreme disparity between the 80% disadvantage of chromosome homozygotes versus the 15% or so attributed to egg-to-adult inviability. There could, however, be many extra fitness components in the population cage. Chief amongst these is fertility. In the female, this is manifested by fecundity, basically number of eggs laid. In the male, the component is presumably competition for mating.

The obvious question to be asked is whether each of these components can be estimated separately, and then combined to see if it accounts for the whole 80%. Although this seems like a straightforward question, in practice it turns out to be anything but.

There was a huge amount of work in the field of fitness components in Drosophila in the period of 1960s to 1980s, much of it reviewed by Simmons and Crow [9] and by Haymer and Hartl [2]. My interest was always restricted to BE, and much of the work on fitness components could not be directly related to this. However Haymer and Hartl reported that the BE method of calculating fitness was an outlier, in the sense that it gave more extreme fitness estimates than other estimation methods. Some time later my colleague Barrie Latter [4] concluded that much of this discrepancy was due to the way that the calculations had been done, and that the fitness estimates from BE were more extreme, but not excessively so, compared to other methods.

2.5. **Continuous populations.**

Although the components can account for much of the BE fitness estimates, the calculations are complex. Partly this is because components such as viability and fertility, at least in their simplest form, can only be combined into a model with discrete, or non-overlapping, generations. However the population cage, and most wild populations, have a continuous generation structure. Development rate, and potentially also time-dependent viability and fertility, come into the equation.

Much theory relevant to this has been developed by Brian Charlesworth [1] and others. Leigh Pascoe, who did a PhD on this topic, made a heroic effort to apply this and additional theory he developed himself to the case of the BE technique. He carried out multiple experiments, including measurements of viability and fertility under competitive conditions. He concluded that competitive effects on viability were more important than my initial studies reported above would have suggested. Surprisingly, however, this competition referred to the way the parents of the cross were raised rather than to competition between the $Cy/+$ and $+/+$ offspring. In retrospect it is probably not surprising that maternal contributions to the egg should confer fitness benefits to the offspring.

Leigh's result seems to have resolved the problem of negative frequency estimates eluded to earlier. But once the possibility of maternal effects, and crowding on such maternal effects, has been admitted, the overall modeling in terms of fitness components becomes almost impossibly complex. What if, for example, the relative fertility of $+/+$ females is affected by whether their mothers were $Cy/+$ or $+/+$, and whether raised under competitive conditions? It is a pity that Leigh's thesis work was never published, It raised as many questions as it solved, but I should have helped or pushed more for publication.

2.6. **Potential BE experiments.**

As is perhaps clear from these remarks, I find the topic of fitness components a little depressing in its complexity, and I wonder whether we will ever really know the full story for even an organism as simple as Drosophila. This is without taking into account the potential effects of 'hybrid dysgenesis', a topic discussed more fully in other places in this PIFFLE, and one that can have important effects when balancer chromosomes are crossed with wild flies.

Regarding the complication of age structure, there was a reasonably simple experiment we ought to have done. This was a careful comparison of the same chromosomes tested by BE in population cages, with their continuous generations, compared to population bottles tipped over every 2-3 weeks to enforce a discrete generation structure. In this way we could have determined how important age-structure was in the production of overall fitness.

I suspect that such experiments are not of interest in the current molecular era. Perhaps they will become of interest again.

I should also record here an experiment that I thought about years ago, but again never got around to seriously looking at. Although the population cage provides what seems like a competitive environment that should emphasise different fitness components, this may be far from the truth. What does a fly do in real life? It has to find food, shelter, mates etc. None of these things are a great problem in the population cage. So might not the fitness differential be even greater if the flies are exposed to real life conditions for which they are actually selected?

One possibility is to set up really large population cages that mimic a wild environment. The experiment that always appealed was to find some small islands in an area where conditions were too harsh to support over-wintering. Then the flies could be introduced in spring, and sampled towards the end of summer, by which time numbers would hopefully have settled down. Probably an unrealistic experiment, especially if it involved intermediate sampling, but a nice thought experiment.

3. OVERDOMINANCE AND DOMINANCE

The experiments on inbreeding depression were motivated by the desire to test the heterozygote advantage hypothesis. As related earlier, the hypothesis proposes that the large amount of variability found in natural populations is not accidental but is selected for.

A problem with the inbreeding depression test for heterozygote advantage is that it has long been known that inbreeding depression doesn't necessarily show heterozygote advantage. There is an alternative, highly plausible, explanation. If heterozygotes are not better than the best homozygote, simply better than the mean of the homozygotes, inbreeding depression is expected.

Three possible selection models are as follows:

Genotype:	A/A	A/a	a/a
-			
Fitness(1) :	1-t	1	1-s
Fitness(2) :	1	1	1-s
Fitness(3) :	1	1-hs	1-s

The first of these models is overdominance or heterozygote advantage, assuming $s, t > 0$. Selection holds the A and a alleles polymorphic in the population in this case. The second and third are both cases where the a allele is deleterious and will be eliminated unless held in the population by mutation. Only model (2) involves true dominance, but if there is some dominance, ie. h must be less than 0.5, model (3) can also be considered in this category.

Models (2) and (3) are not controversial. There are undoubtedly many cases of deleterious genes at some sort of mutation-selection balance. Heterozygote advantage or overdominance, on the other hand, is a hypothesis that needs further substantiation. Although there are a small number of known cases, the notion of widespread heterozygote advantage maintaining variability is one that needs further substantiation.

What is the importance of this distinction? Lewontin in his influential 1974 book [5] summarised two world views about the nature of variability in populations. Under the 'classical hypothesis', the variability seen in natural populations is basically just departure from some ideal genotype, brought about by mutation. Under the 'balance hypothesis', on the other hand, most or all of the variability found in natural populations is there for a selective reason. At a genetical level that can be modeled in a simple way, these two world views correspond to the dominance and overdominance models respectively. It is, however, asking a lot from the heterozygote advantage model to explain the complexity of variability found in surveys such as Hapmap for human populations [16] [17].

3.1. **Inbreeding consequences.**

In interpreting inbreeding data, all three hypotheses predict inbreeding depression. The levels of inbreeding depression are somewhat different for the three models. Considering a single locus, if the frequency of the a gene is q, the average amounts of inbreeding depression brought

about by complete homozygosity of a chromosome are:
Model (1) $(s + t)q(1 - q)$
Model (2) $sq(1 - q)$
Model (3) $s(1 - 2h)q(1 - q)$

All three involve the frequency term $q(1 - q)$ multiplied by a selective value term. Note that the third term will only be positive if there is some dominance, ie $h < 0.5$. The major difference between the three is, however, that values of q are expected to be very small for the case of mutation-selection balance, particularly for model (3). If equilibrium is assumed, due by selection in model (1) and mutation-selection balance in models (2) and (3), the expected inbreeding depressions are approximately:
Model (1) $st/(s + t)$
Model (2) $\sqrt{us}$
Model (3) $u/(2h)$
Potentially the value under model (1) is greater than for the other two, given that the latter are functions of the mutation rate.

What is unknown is how many loci of the three different types there are, and how the selective values at different loci cumulate. If there are many more loci of type (2) and (3) than of type (1) then dominance may play a larger role in inbreeding depression than overdominance.

3.2. **The di-chromosomal cage - a test for dominance.**

We [21] developed a test that should detect deleterious recessive genes, models (2) and (3). The idea was to make up population cages heterozygous for just two wild type chromosomes. In these populations all deleterious recessive genes should be present at a frequency of 50%. In subsequent generations of these populations, these genes should be selected against, or *purged*, leading to chromosomes of higher homozygous fitness. The experiment is a rather long and complicated one, and its design is shown in Figure 5.

Note that the cages shown with white tops are di-chromosomal cages containing no balancer chromosomes. In order to do fitness measurements, it is necessary to set up BE cages, in which single chromosomes are sampled according to the principles of Figure 1 and placed into cages, shown with black tops.This sampling of chromosomes is done at the beginning of the experiment, and then at certain intervals, eg. 6 months and one year.

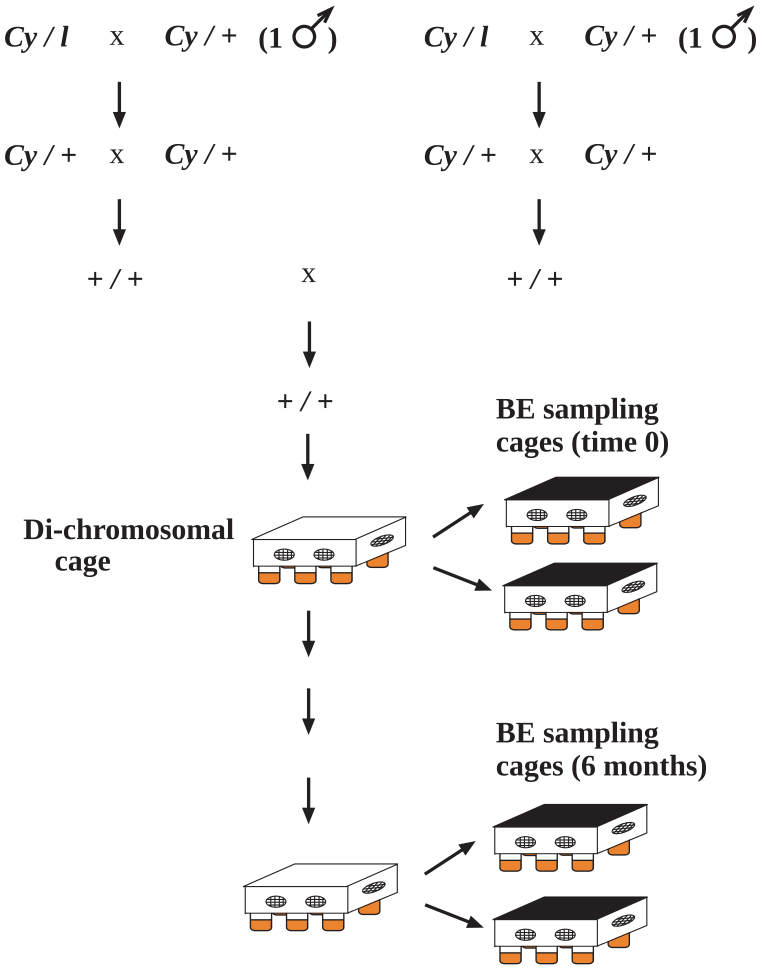

Figure 5. The di-chromosomal population experiment

We set up several di-chromosomal cages, in each case taking care to choose wild type chromosomes sampled from different populations. These populations in our case were different wineries from the Hunter Valley district of NSW in Australia. Tyrrell's winery, for those who know it, was always a favourite, but there were several where we could regularly sample 'wild type' flies. The choice of two chromosomes from different populations almost guarantees that any low-frequency deleterious allele will be at 50% frequency.

Deleterious alleles starting at 50% frequency should be under high selective pressure. In theory, the selection against a deleterious recessive allele is maximised when the frequency is $\frac{2}{3}$, but the selection against an allele at frequency $\frac{1}{2}$ should still be high, certainly compared to the frequency when q is low and the allele is rarely present in homozygous condition. The net result should therefore be a purging of deleterious genes, as shown by increased homozygous chromosome fitness.

Unfortunately it is not possible to write down the expected rate of purging of deleterious genes, because this depends on the value of s, the selective disadvantage. Extreme possibilities for the deleterious allele hypothesis are that there are few genes each of which has a high value of s, or that there are many such genes each of which has low s. The lower the number of such genes, and the higher the value of s, the more effective the di-chromosomal experiment should be at purging deleterious alleles.

3.3. **Negative results from the di-chromosomal experiment.**

The results of the experiment were entirely negative [21]. We never found anything like a high fitness homozygous chromosome. There was no systematic change in mean fitness over the course of the experiment. Although from the point of view of the heterozygote advantage hypothesis a negative result was preferable, this was a disappointing result from the point of view that a positive result is always more stimulating than a negative result. We were now in a position of having to look for reasons why the experiment didn't produce positive results.

One possible reason is the one mentioned above, that the more deleterious genes there are, and the smaller the selective disadvantage of each, the smaller the expected rate of loss of such genes. There is a limit to this argument, however. There can't be an infinite number of loci on a chromosome, so the expected number of deleterious alleles at such loci can't, in turn, be all that great. We did simulations to try to see how many loci there need to be before the rate of loss becomes undetectable, and this indicated that we ought to have seen it.

A second reason is that the design of the experiment guarantees that there will be 100% LD at the start. In the absence of recombination the frequencies will never change. Again we relied on simulation to show that over the course of 6-12 months, estimated as 9-18 generations, there should be enough recombination to overcome this initial LD.

A third possibility is that new deleterious genes by mutations are produced at a rate sufficiently high to balance the loss of old ones by selection. In theory this explanation ought not to work either. The frequency of mutations is at a mutation - selection balance. We are increasing the frequency of selection enormously by bring frequencies to 50%. Therefore this balance should be upset.

There is, however, a caveat to the mutation argument. The crosses of Figure 5 are potentially those that give rise to 'hybrid dysgenesis'. There is a whole section of this PIFFLE devoted to this topic - the point is that mutation rates can be increased by such dysgenesis. We made the crosses in a way that should have reduced this effect as much as possible. But we were never sure that we had totally eliminated it. That was the major reason for the next experiment.

3.4. The di-parental experiment.

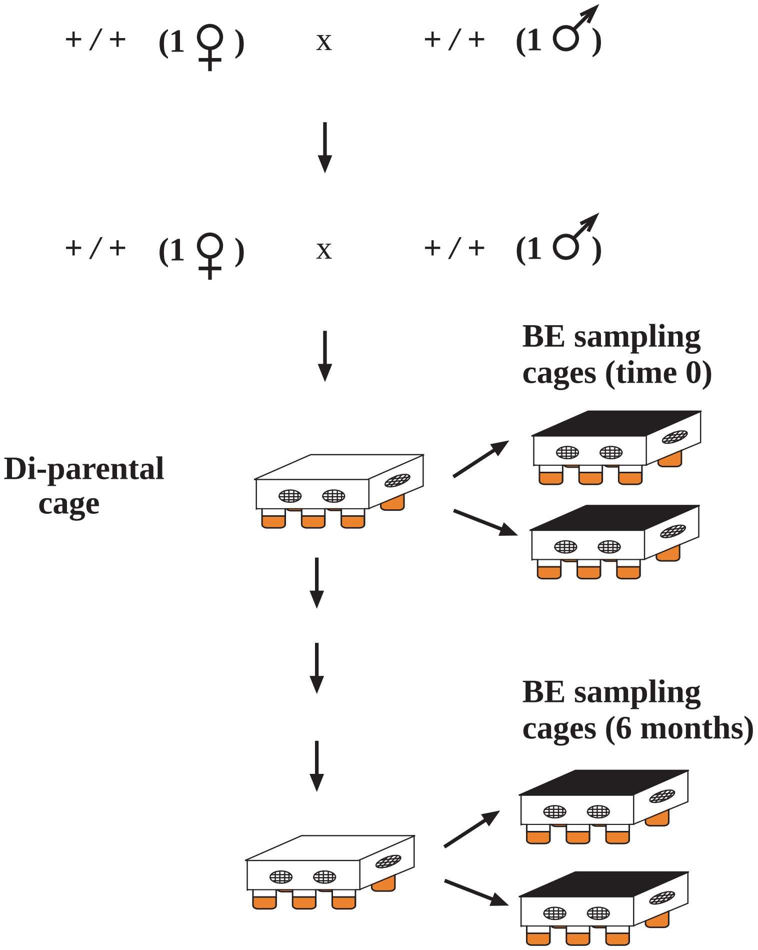

Figure 6. The di-parental population experiment

Figure 6 shows the very simple crosses for this experiment. A single male and single female from two different populations are crossed. This is followed by a second cross of a single female and male, and the progeny are used to found the di-parental population.

The actual procedures that we used were somewhat more complicated. We didn't particularly want to start cages with either a large number of lethals or a large number of chromosomal inversions which would inhibit crossingover. So we had to start with many of crosses of the type shown in Figure 6 and retrospectively discard populations where lethals and inversions were found.

The theory behind the experiment is shown in Figure 7. Any deleterious gene, assumed to be present just once in the first cross, has a 1/4 chance of being absent from the di-parental cage, a 1/2 chance of being present once and a 1/4 chance of being present twice. Thus there is a 2/3 chance that any deleterious gene in the population will start at 25% frequency, and a 1/3 chance that it will start at 50% frequency.

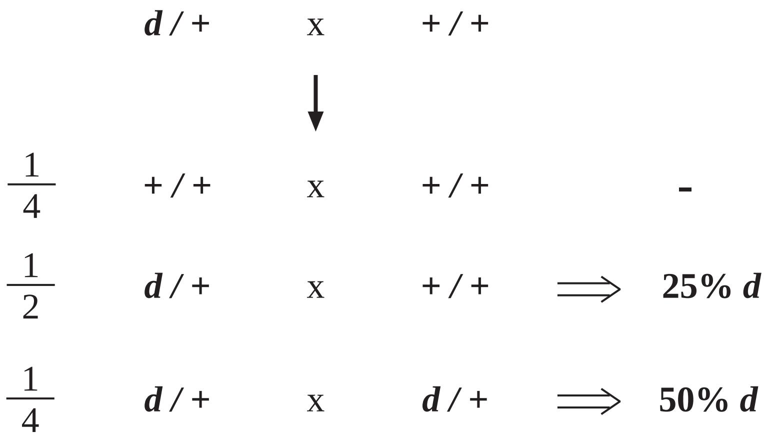

Figure 7. Frequencies of deleterious genes in the di-parental cage

The numbers are not quite as favourable as for the di-chromosomal cage, where all deleterious genes start at 50%. However simulation shows, surprisingly, that the purging of deleterious genes is no less rapid. In fact depending on the value of s it could be more rapid. The reason for this result is that the initial LD is not as high in the di-parental population as it is in the di-chromosomal population, which improves the efficiency of the procedure.

The great advantage of the procedure is that no crosses to laboratory strains are involved right up to the time fitness measurements are made. There is therefore no possibility of unwanted hybrid mutation effects coming into the procedure.

3.5. **Another negative result.**

Once again, the experiments gave no sign of an increase in homozygous fitness. These experiments have not (yet) been written up for publication. Any day now.

So what is to be made of another apparent failure to purge deleterious genes? In theory it means that the dominance hypothesis for explaining inbreeding depression or chromosomal heterosis has to be modified in some way. But I still can't bring myself to accept the alternative that all of the chromosomal heterosis is explained by heterozygote advantage. Somehow this seems too simple, or glib, a hypothesis to be able to explain the maintenance of variability in the population. I've talked about Associative Overdominance in another section of this PIFFLE, and perhaps one needs to take this a little more seriously.

4. BACK TO VIABILITY

The negative results in my original competition experiments came as a surprise to me. If a particular +/+ homozygote had egg-to-adult viability of 90% compared to $Cy/+$, it seemed that this 90% held true whether there were few or many larvae competing for the available food. If that is the case, what is causing the additional 10% mortality among the +/+ genotypes? Obviously there might be some developmental stage affected. But if this is the case, why do 90% survive this step and 10% not? All of the genotypes are presumably the same for whatever developmental stage is involved.

Perhaps this is a philosophically naive way of looking at the problem. *A posteriori* any small difference can be responsible for an all-or-none effect. Anyway it seemed to me that one possible solution to this quandary of what differs between the flies that live and those that die might be genes on other chromosomes. So I set out to test this hypothesis.

4.1. **An experiment to select for increased viability.**

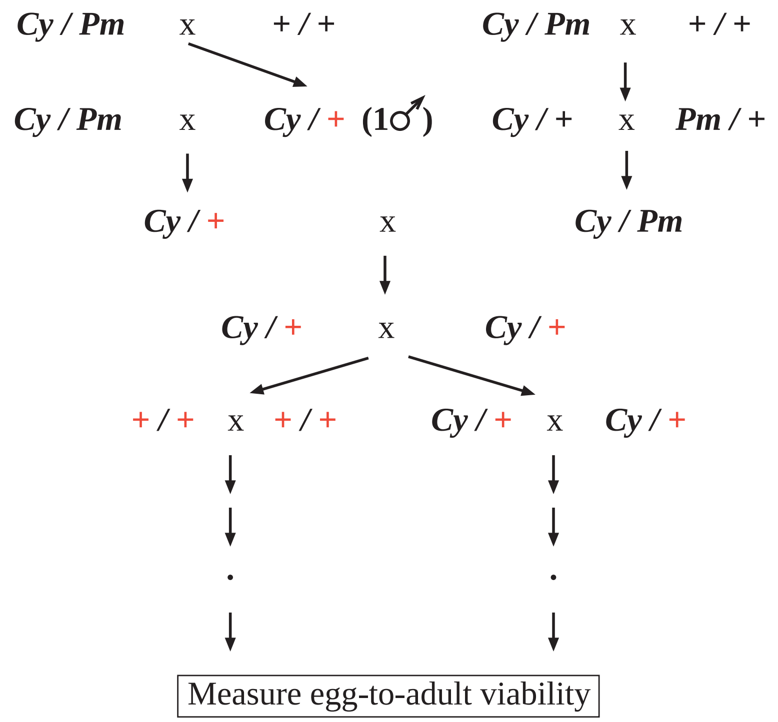

Figure 8. Experiment to select for increased viability

Figure 8 shows the procedure. Several things need to be noted. First, the '*l*' symbol has been replaced by *Pm*. The principle is the same, except that the experiment requires that this lethal-containing chromosome be recognisable. *Pm* is an eye colour marker, Plum.

The second complication, the crosses on the right of the first three lines, ensures that inbreeding only occurs on the second chromosome. While only the second chromosome is shown, the fact that multiple parents are used ensures that the genotypes for chromosomes other than the second, the genetic background, are mixed. The + second chromosome, shown in red in the figure, necessarily comes from a single original copy.

The purpose of the experiment is to select for genes that increase viability in the $+/+$ genotype. As argued above, such genes can potentially be on any chromosome other than the second. The series of crosses on the left was carried on for 25 generations before crossing to Cy/Pm to measure viability in the usual way.

The crosses on the right are the control crosses, in which the $+/+$ genotype never occurs. In theory, therefore, the viability should stay much the same. The parents in this case need to be selected in each generation, rather than just tipped over as in the crosses on the left.

This experiment was so successful it was written up and accepted by Nature [12]. Unfortunately the reason for this had nothing to do with the original purpose of the experiment. The Nature paper concentrated primarily on the control crosses. What happened here is that the +/+ genotype became lethal over the course of the experiment, independently of the genetic background. This was totally unexpected, showing that mutation must be occurring at a high rate. There was no obvious reason for this. In those days Nature (sometimes) accepted phenomenological papers such as this even if there was no molecular explanation. The great evolutionary biologist John Maynard Smith once confessed to me after a few drinks that he was a reviewer of the paper, maybe the only one, and recommended its acceptance, although I suspect he had misgivings.

The end result of all of this was that I lost interest in fitness experiments and went over to experiments to try to work out what was responsible for the high mutation phenomenon. This work, which lasted for 30 years or more, is all written up in the chapter entitled 'Hybrid dysgenesis and P elements'.

4.2. **Back to the results.**

Looking back at the non-control results in the Nature paper, the experiment was not a total failure. The results were as follows:

Chromosome	Viability before	Viability after
1	0.91	0.89
2	0.90	0.95
3	0.64	0.88
4	0.37	0.83
5	0.33	0.53
6	0.31	0.88

Although the viability has not increased to unity, nevertheless there is an improvement in all but No 1, the highest viability original chromosome. Each of the readings above is based on two replicates, and the agreement between replicates was not bad. So it looks as if genetic background does control some part of the life-death decision.

Clearly there is room for additional experiments along these lines. A proper control needs to be done now that the hybrid mutation effect is understood. And there should be more chromosomes tested at different time intervals. Ultimately this is a developmental biology question and needs to be understood in these terms.

4.3. **Choice of balancer chromosome.**

One last question that I find intriguing in the viability experiments is the choice of balancer chromosome. As mentioned previously there are many such balancer chromosomes. Some of these, such as $Cy(SM1)$, appear to have no effect on viability. Others such as $Cy(SM5)$, which is reputed to be a better suppressor of all recombination, survive at a rate closer to 50%.

The question is, which of these should be used in a viability experiment such as in Figure 1? I recall, without having any reference, that Dobzhansky and Bruce Wallace, two of the chief proponents of the technique, recommended the use of high viability balancer chromosomes. It is a recommendation I always believed in.

The alternative possibility is that a low viability balancer chromosome is more 'sensitive' in detecting changes in the test genotype. This ought to be the case only if viability involves some sort of competition between genotypes. Our experiments described earlier showed little indication of such competition.

Nevertheless one indication of potential increases in sensitivity of low viability balancer chromosomes comes from an experiment of Trudy Mackay and colleagues [6]. They studied the effect of single P element insertions on viability and found high effects on homozygous viability, of the order of 10% per insert. We used the BE technique to measure homozygous fitness, currently unpublished, and were unable to detect effects on overall fitness. In theory BE should be more sensitive. The possibility, therefore, is that the low fitness balancer chromosomes used in [6] might show up viability differences more sensitively. It would be of interest to study the use of different balancers in more detail.

5. LOOKING BACK

In an era where so much is known at the molecular level it seems striking that there are such gaps in knowledge about how selection works. As argued above, we don't really know whether the dominance model or the overdominance model provides a more accurate description of the average locus.

Perhaps laboratory fitness experiments will become fashionable again. They are at least cheap compared to molecular experiments.

References

[1] B. Charlesworth. *Evolution in Age Structured Populations.* Cambridge U.P., 1980.

[2] DS Haymer and DL Hartl. The experimental assessment of fitness in drosophila. i. comparative measures of competitive reproductive success. *Genetics*, 102(3):455–466, 1982 Nov.

[3] J L King. Continuously distributed factors affecting fitness. *Genetics*, 55:483–492, 1967.

[4] B. D. Latter and J. A. Sved. A reevaluation of data from competitive tests shows high levels of heterosis in drosophila melanogaster. *Genetics*, 137(2), 1994.

[5] R. C. Lewontin. *The genetic basis of evolutionary change.* Columbia U.P., N.Y., 1974.

[6] R F Lyman, F Lawrence, S V Nuzhdin, and T F Mackay. Effects of single p-element insertions on bristle number and viability in drosophila melanogaster. *Genetics*, 143(1):277–292, 1996 May.

[7] RD Milkman. Heterosis as a major cause of heterozygosity in nature. *Genetics*, 55:493–495, 1967.

[8] C A Mourao, F J Ayala, and W W Anderson. Darwinian fitness and adaptedness in experimental populations of drosophilia willistoni. *Genetica*, 43(4):552–574, 1972.

[9] M J Simmons and J F Crow. Mutations affecting fitness in drosophila populations. *Annu Rev Genet*, 11:49–78, 1977.

[10] D Sperlich and A Karlik. The genetic conditions in heterozygous and homozygous populations of drosophila. i. the fate of alien chromosomes. *Genetica*, 41(2):265–304, 1970.

[11] J. A. Sved. An estimate of heterosis in drosophila melanogaster. *Genet Res*, 18(1), 1971.

[12] J. A. Sved. Short term heritable changes affecting viability in drosophila melanogaster. *Nature*, 241(5390), 1973.

[13] J. A. Sved. Fitness of third chromosome homozygotes in drosophila melanogaster. *Genet Res*, 25(2), 1975.

[14] J. A. Sved and F. J. Ayala. A population cage test for heterosis in drosophila pseudoobscura. *Genetics*, 66(1), 1970.

[15] J. A. Sved, T. E. Reed, and W. F. Bodmer. The number of balanced polymorphisms that can be maintained in a natural population. *Genetics*, 55:469–481, 1967.

[16] The-International-Hapmap-Consortium. A haplotype map of the human genome. *Nature*, 437:1299–1320, 2005.

[17] The-International-Hapmap-Consortium. A second generation human haplotype map of over 3.1 million snps. *Nature*, 449:851–861, 2007.

[18] M L Tracey and F J Ayala. Genetic load in natural populations: is it compatible with the hypothesis that many polymorphisms are maintained by natural selection? *Genetics*, 77(3):569–589, 1974 Jul.

[19] B. Wallace. *Genetic load, its biological and conceptual aspects.* Prentice-Hall, NJ, 1970.

[20] A. N. Wilton and J. A. Sved. X-chromosomal heterosis in drosophila melanogaster. *Genet Res*, 34(3), 1979.

[21] AN Wilton, MG Joseph, and JA Sved. Can chromosomal heterosis in drosophila be explained by deleterious recessive genes - negative results from a dichromosomal population test. *Genetical Research*, 53:129–140, 1989.